2004

Development of a machine vision system for corn plant population, spacing and height measurement

Dev Sagar Shrestha
Iowa State University

Development of a machine vision system for corn plant population, spacing and height measurement

by

Dev Sagar Shrestha

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Agricultural Engineering (Agricultural Power and Machinery)

Program of Study Committee:
Brian L. Steward, Major Professor
Carl Bern
Jennifer Davidson
Thomas C. Kaspar
Stuart Birrell

Iowa State University

Ames, Iowa

2004

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Dev Sagar Shrestha

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

THIS WORK IS DEDICATED TO ALL THOSE WHO NEED FOOD ON THE TABLE TO FEED THEIR CHILDREN.

TABLE OF CONTENTS

Acknowledgements

I would like to extend my sincere gratitude to my advisor Dr. Brian L. Steward for his continuous guidance, inspiration and leadership throughout my PhD program. I believe that this work makes a foundation for the development of a complete sensing system that would be useful for the application of site specific management.

Thanks to Prof. Carl Bern for mentoring and giving valuable inputs to make this work more practical. Thanks to Dr. Stuart Birrell for his help and guidance with statistical analysis and other valuable suggestions. Thanks Prof. Thomas Kaspar for enriching my knowledge for the agronomic importance of this research. Thanks Dr. Jennifer Davidson for her valuable comments specifically in stereo vision and image processing in general.

My thanks go to administrative staffs Mary Ellan, Judy Reschauer, Sue Ziegenbusch and Barbara Kalsem for their continual support and help. Thanks to Richard VanDePol for providing necessary support and facility in the farm. Thanks to undergraduate students Cory Van Wyngarden, Eric Stephan, Brian Crawford, Nick Krueger, Sandra Wenke, Rudy and Matt for programming and construction support. Thanks to Mitch Miller, Sadettim Yildirim and Mark Westphalen, my office mates, for their help and support at times.

Finally thanks to my parents Ganga Sagar and Bikumaya, brothers Som and Krishna Shrestha, my wife Sinora and son Amesh for their moral support to complete my PhD with great enthusiasm. Actually it was my eldest brother Som Sagar who set a high standard of education for me to accomplish in my very childhood. I feel truly blessed to have so many people help and mentor me to accomplish something in life. I hope with God's help I will be able to fulfill the expectations of my mentors to serve humanity in a meaningful way.

Abstract

Precision agriculture has a great potential to increase yield by optimizing agricultural inputs at a sub-field level through improved spatial management. Extensive knowledge of agricultural inputs and output is required to effectively use precision agriculture technology to have an economic benefit to producers. Two field variables that have a great impact in the final yield are plant population density and inter-plant spacing.

In this research, a system was developed to measure the spatial variability of early stage plant population density, spacing and plant height. The Truncated Ellipsoidal (TE) method was developed to segment plants from background. It was found that this method had the lowest amount of classification error compared to NDI and Bayesian methods in different daylight conditions. TE was used to segment plant from background throughout this research.

A patch matching algorithm was developed to sequence for video frames of corn row videos. A statistical criterion from Chebyshev's theorem was introduced in the algorithm to make frame matching more robust. This algorithm reliably sequenced the video frames of corn row scenes acquired by a commercial digital video camera on a vehicle traveling at 1 to 2 m/s.

Two features were extracted from the segmented and mosaicked images and second order statistics of these features were calculated. Otsu's method was modified to compensate for noise present and was used to calculate the optimum threshold to distinguish between plant and non plant regions. Two additional features from vegetation regions were extracted and used to differentiate between weeds and the corn plants. Algorithm performance was analyzed across three tillage treatments, three growth stages from V3 to V8, and three population densities varying from 27,000 to 81,500 plants/ha. Overall, the algorithm estimated the number of plants in 6.1 m crop row lengths with an RMSE of 2.1 plants. The mean measurement error was significantly different across tillage treatments, but no evidence of significant differences was found across growth stages and plant populations. The error variance at V7-V8 growth stages was significantly higher than that at V3-V4 growth stages. No significant differences were found between mean measured and estimated plant spacing distances.

For implementation of this plant density measurement and spacing measurement algorithm, a component-based software architecture was developed to automate site specific field data acquisition, processing, and geo-referenced plant parameter extraction. The architecture supported acquisition and processing of different data streams such as digital video or digital serial communications. A key component of this architecture was a supervisor class, which communicated and coordinated the operations of all other modules and classes. Standard software components were used in the architecture to do data acquisition. Based on this framework, early stage corn population estimation (ESCOPE) software was developed which grabbed pre-recorded digital video from a vehicle-mounted camera that was passed over corn rows and acquired GPS-NMEA strings which were modulated and recorded on the audio channel. Reusability and extensibility characteristics were demonstrated by adding a class to acquire images from the hard drive and also by deriving a new image analyzer class to extract an additional feature. The architecture forms a general framework for developing reusable and extensible software for field data sensing systems.

For the crop height measurement, two different sensing approaches, stereo vision and ultrasonic, were investigated as candidate technologies for vehicle-based corn height sensors. For the stereo vision method, a chain code-based stereo correspondence technique was developed to determine the disparity in the stereo image pair. Images were taken using one camera from a series of precisely controlled locations to generate the stereo effect. The ultrasonic sensor measured the distance to an object by detecting the time of flight of ultrasonic sound waves. The echoes from the plant canopy were recorded, and collar height of the plant was estimated. A good correlation was found between the measured and estimated height using both stereo vision and the ultrasonic sensor. For the stereo vision sensor, r^2 between the maximum plant height and estimated height was 0.76. For the ultrasonic sensor, r^2 between the 25^{th} percentile of the group height statistics and plant collar height was 0.75.

Chapter 1. Introduction

Spatial variation

The National Research Council (1998) refers to precision agriculture (PA) as a management strategy that uses information technologies to bring data from multiple sources to bear on decisions associated with crop production. The concept of site specific management (SSM) was first initiated for application of fertilizers to Midwestern corn and soybeans; it has been adapted to a variety of practices, crops, and countries. From variable rate of fertilizers, it now encompasses all management practices on a spatial and temporal basis (Robert, 2000).

The key idea behind PA is to measure and manage spatial variability to optimize the crop production system. Spatial variability can be categorized into yield, field, soil, crop, management, and variabilities due to anomalous factors. These variables can also be classified as controllable variables and variables that can not be controlled. For example irrigation water, seed type, amount of fertilizer and pesticide to be used are controllable variables where as amount of sunshine, soil type, amount and time of rain, and temperature changes are uncontrollable variables or at least not in direct control of a farmer. The objective of good field management should be to manage controllable factors to match with the uncontrollable factors to optimize the overall crop production. According to Liebig's law of minimum, "crop yield is a proportional function to the scarcest nutrient available to the plant" (Berck et al., 2000). That means if one of the nutritive elements is deficient or lacking, plant growth will be poor even when all the other elements are abundant. Any deficiency of a nutrient, no matter how small an amount is needed, will hold back plant growth. If the deficient element is supplied, growth will be increased up to a point where the supply of that element is no longer the limiting factor. Increasing the supply beyond this point is not helpful, as some other elements would then be in minimum supply and become the limiting factor. Liebig's law of minimum assumes that the input factors are irreplaceable and does not account for the crop adaptability. For instance, according to Liebig's law, if water availability is limiting, no amount of nutrient input can increase the yield. There are certain modification proposed by other researchers (Berck and Helfand, 1990; Sinclair and Park, 1993), however,

Liebig's law of minimum still can be used to show the importance of optimizing all input levels. Because of the fact that these uncontrollable variables vary from one point to another in the field, we should vary other inputs such that no applied inputs are wasted.

In addition to the variables listed above, there are other uncontrollable variables associated with controllable variables which are difficult to avoid. For example, when the seeds are placed, there will always be some tolerance in the precision of the seed placement and there is random variation in seed germination. These variations can be minimized by continuous improvement of equipment design and careful operation, but it is difficult to eliminate these variations altogether. This random variability associated with assumed controllable variables may offset yield prediction. Random variation in seed placement and germination cause the plant population to deviate from target plant population and interplant spacing. In parts of a field, where the plant population density is higher, the demand for an agricultural input will be high in contrast to the areas where the plant population density is lower.

Challenges of precision agriculture

The relationship between plant population density and amount of agricultural inputs required may not be linear relationship. This nonlinearity comes from interactions of other factors such as sunlight interception and water availability. Studies on the effects of nitrogen uptake, plant population density, and yield confirmed this nonlinear relationship between nitrogen uptake and plant density (Kayode and Agboola, 1981; Mathers and Stewart, 1982). This nonlinear relation makes optimization of different input levels difficult. Quantifying the amount of all nutrients necessary for plant growth is neither feasible nor economical at this point. Even if we were able to quantify the naturally occurring elements for individual plants, the amount of other agricultural inputs to produce optimum yield would be still unknown.

Because of these complexities, PA is still much in its infancy. PA is a holistic approach, but today only some the elements of this whole system have been studied to correlate with yield. Robert (2000) pointed out that further research is needed to make the PA more successful in the areas including real time sensing for soil and plant characteristics. He further concluded that it would take time before PA is adopted widely; however he referred to PA as the agricultural system of 21st century.

Machine vision based sensors

There have been studies in the area of precision agriculture to relate agricultural inputs to the crop production (Strock and Malzer, 2000; Ward and Cox, 2000). Sensors have been developed to measure some of the soil and plant characteristics. However, there are other variables for which practical sensing technology have not been developed. Some field variables could be measured indirectly using an indicator variable. An indicator variable is a variable that reflects the status of the original variable but is easier to measure and quantify than the original variable. For example, leaf color could be used as an indicator variable for nitrogen deficiency. When it comes to indirect measurement of input variables, machine vision can be used to detect and measure many important field variables. Machine vision can be used for non destructive measurement of important field characteristics like plant health (Hetzroni and Miles, 1992), weed infestations (El-Faki et al., 2000; Tang et al., 2000), plant shape and size (Nishiwaki et al., 2001). This makes machine vision one of the best options to develop for field variability measurement. This also obviates the necessity to have separate sensors for each of the field variability being measured. One difficulty in using machine vision, however, is the need for large amounts of storage media and processing capability. In the past, with smaller storage media capacity and slower processor speed, it was almost impossible to measure the field variability using machine vision. The advent of compact but high capacity digital storage devices like MiniDV® tape, computer speeds up to 3.2 GHz for commercially available personal computers and rapid growth in processing speed have made it possible machine vision to use to measure field variables.

As it has been pointed out in previous sections, accounting for all the field variables is almost an impossible task. There are still many unknown factors affecting crop yield. Therefore, as a solution to relate inputs to the final yield, it would be best to start with the variable that can be measured easily and have the known impact on crop yield. Once the relationship has been established, then more field variables can be incorporated into the model until this addition is no more economically or environmentally beneficial.

Two such variables that have major impacts on corn yield are crop plant population and spacing. Machine vision can be used to directly measure the plant density and spacing. Another overall indicator of input adequacy is plant growth rate. If a plant is supplied with an

adequate amount of nutrients and there are no limiting factors, the growth rate of the plant should be 100% of its potential. Whenever there are some limiting factors, it is likely that the growth rate of the plant will be reduced. The growth rate of the plant can be obtained by periodic measurement of plant height in the field. Machine vision can also be used to measure the plant height. Growth rate could be used as an indirect measure of many uncontrollable and controllable inputs. As one of the objectives in precision farming is to quantify those natural resources available to the plant to supply the deficit, growth rate measurement could be a good indicator variable to collectively quantify those factors.

Important field variables

Plant populations that are higher and lower than the optimal plant population can reduce the yield. Duncan (1958) found that the grain yield was maximum at a certain plant population depending upon variety and other characteristics. Wiley and Heath (1970) investigated the relationship established by different researchers between plant population density and crop yield and found the predictions had similar trends of yield maximization at a certain population density.

In addition to the production side, sensing plant population will enable evaluation of the planter performance and seed germination rate. For seed germination, it is not only the question of whether a seed will germinate or not, but for PA, it is equally important how long it takes for plant emergence. If the seed is not germinated in due time, it may have serious consequences in yield. Especially when population density is high, a plant has to compete for sunlight and other essential elements. Plants germinating earlier will have a competitive advantage and will tend to have more vigorous growth. Plants germinating later may already be in the shade of previously germinated plants. Having a smaller root system than a taller plant, it will be less competitive for nutrients and water and hence will be weaker. Weaker plants are more susceptible to the disease and pests. Therefore, early germinated plants will be healthier and will tend to give more yield than the plants germinated later. With so many interdependent variables, even a slight variation in one variable can make a big difference in yield. Plant population determination and growth rate measurements can quantify this kind of variability in the field so that the management can take the right decision to reduce the economic loss.

While quantifying the availability of nutrients and other inputs, it is important to consider the plant spacing distribution as well or to make comparisons on a per-plant basis. It is not only important to know how much was applied but also to know how much was available to the plant. If plants are clustered in one location and dispersed in another, the nutrients and sunlight may be wasted in the areas of sparse distribution. Therefore, even if all other variables are identical in two different parts of the field, mere variation in plant spacing makes the performance of plants grown with different population density be different. Therefore it is imperative to measure the spatial distribution of plant position along with other variables for true understanding of the effect of other variables on individual plants.

Need for a complete system

If plant population density could be estimated automatically at its early stage of growth, early yield potential estimates could be made based on the number of plant available in the field to produce grain. Automated plant population estimation could also enable the evaluation of other factors like seed germination rate and land fertility variation within a field. In addition, if plant population data of subsequent years could be recorded along with other factors like moisture content, amount of precipitation, temperature, the effect of different variables determining the final yield of crop could be measured. Manual counting and recording of population density is a tedious task and subject to error. In addition, it would not be feasible to manually count a large field area.

With a system that will measure both plant population density, spacing and plant height not only could the amount of natural inputs available for plant growth be quantified, but also depending on availability, future decisions could be made about the amount of other agricultural inputs to apply. This type of site specific management will help to optimize the field production at a sub-field level, so that optimum productivity could be maintained in each sub section of a field.

Therefore in this research, a machine vision based approach for determining plant population density and height was developed. Machine vision has potential for future proliferation of its use. For example in the future other variables like plant canopy color, soil color, soil texture, and weed density could be measured using the same system.

Objectives

The goal of this research was to develop a complete machine vision based sensing technology for measuring early stage corn plant population, spacing and height. Such a system should be capable of acquiring various types of field data and processing. In order to accomplish this goal following objectives were set:

1. Develop a robust segmentation system to separate plant from the background for varying lighting conditions in the field.
2. Develop a machine vision algorithm to locate and count the individual plants from the segmented image.
3. Develop an image sequence mosaicing procedure for video analysis.
4. Develop general software architecture to implement field data collection, processing and information extraction for precision agriculture use.
5. Develop a robust correspondence method for height determination.
6. Evaluate the performances of developed algorithms.

Dissertation Organization

There are a total of five papers included in this dissertation. The first and second papers have already been published. The third and the fourth paper were submitted for publication and are currently under review. The fifth paper was presented in the ASAE annual meeting conference at Chicago in 2002.

Different plant segmentation techniques are discussed in chapter three. In this paper a truncated ellipsoidal (TE) decision surface was developed to segment plants from background. A feed-forward neural network was trained to adjust the decision surface for various lighting conditions. The segmentation performance of this approach was compared with two different conventional methods of segmentation.

In chapter four, automatic corn plant population measurement using machine vision is discussed. In this chapter, a video frame correspondence algorithm was developed and the TE decision surface was used to segment plant from the background. The TE decision surface was used without a neural network for average lighting conditions for segmenting

plant from background. A digital video camcorder was used to take the video of a corn row. The results were compared with manual measurements for performance evaluation.

Plant spacing measurement and weed-corn plant separation are addressed in chapter five. In this paper the single corn plant, multiple corn plants, weed and background were separated with using statistical approach. Otsu's method was modified to calculate the thresholds for class separation. Algorithm performance was analyzed across three tillage treatments, three growth stages, and three population densities.

Component-based software architecture for field data acquisition, processing, and information extraction was developed and is discussed in chapter six. A key component of this architecture was a supervisor class, which communicated and coordinated the operations of all other modules and classes. Standard software components were used in the architecture to do data acquisition. Based on this framework, early stage corn population estimation (ESCOPE) software was developed which grabbed pre-recorded digital video from a vehicle-mounted camera that was passed over corn rows and acquired GPS-NMEA strings which were modulated and recorded on the audio channel. The plant counting and spacing measurement technique developed in chapter five was implemented to develop and locate the plant and measure the spacing.

In chapter seven, two different sensing approaches, stereo vision and ultrasonic, were investigated as candidate technologies for vehicle-based corn height sensors. For the stereo vision method, a chain code-based stereo correspondence technique was developed to determine the disparity in the stereo image pair. Images were taken using one camera from a series of precisely controlled locations to generate the stereo effect.

Chapter 2. Literature Review

Optimum yield

Total production and yield per plant are two different things. There is always an optimum plant population that will give the maximum yield. Plant populations that are higher and lower than the optimal plant population can reduce the yield. Duncan (1958) found that the grain yield was maximum at a certain plant population depending upon variety and other characteristics. Wiley and Heath (1970) investigated the relationship established by different researchers between plant population density and crop yield and found the predictions had similar trends of yield maximization at a plant population density.

Early corn plant population sensing has the potential to be a useful tool in evaluating spatially varying emergence patterns. It also may find usefulness in evaluating planter performance since yield reduction may occur because of corn spacing variation increases (Nielsen, 1995). Machine vision technology is a promising approach to early corn population sensing. If plant population density could be estimated automatically at its early stage of growth, early yield potential estimates could be made based on the number of plants available in the field to produce grain. Sensors have been developed to measure corn population at harvest (Sudduth et al., 2000; Nichols, 2000). Comparison of early plant population and harvesting time population may be used to estimate the population density required at planting time to achieve desired population density at harvesting time. Automated plant population estimation could also enable the evaluation of other factors like seed germination rate and land fertility variation within a field. In addition, if plant population data of subsequent years could be recorded along with other factors like soil moisture content, amount of precipitation, temperature, the effect of different variables determining the final yield of crop could be measured. Manual counting and recording of population density is a tedious task and subject to error. In addition, it would not be feasible to manually count a large field area.

Sensor need

A major limitation to identifying and mapping yield-limiting factors in fields is the availability of appropriate on-the-go sensing technologies for plant growth (Sadler et al.,

2000). Combine yield monitors are one example of sensing technology that has proven successful in measuring crop spatial variability in real-time with high spatial measurement density. However, yield monitors are relatively inaccurate for small areas (Colvin et al., 1999) and can only measure plant response after the growing season is completed. The capability to measure crop growth responses to spatially varying factors at several times during the growing season would provide the opportunity to identify, map, and manage crop stress to minimize its effect on yield. For example, by making sequential measurements of plant height across a field and calculating a growth rate throughout the crop-growing season, it may be possible to identify when and where stress is occurring and the probable causes of that stress. Sammis et al. (1988) used repeated, manual measurements of plant height as an indicator of water stress, evapotranspiration, and yield of irrigated corn, but they did not attempt to determine spatial patterns. Measuring spatial and temporal patterns of crop growth on a field scale has potential as a diagnostic tool for identifying crop stress or crop responses to spatial variability. In general, however, traditional manual crop growth measurements are too destructive or too labor intensive for sequential crop growth measurements to be used on field scale for either commercial or research applications.

Image-based crop growth measurement has been shown to be effective in measuring and modeling crop plant growth in laboratory or greenhouse applications (Morden et al., 1997; Van Henten and Bontsema, 1995; Tarbell and Reid, 1991 a,b; Tarbell et al., 1991). However, vehicle-based sensing systems have not been developed to make repeated, non-destructive, non-contact crop growth measurement on field scales. A stereo vision system should make these measurements possible.

Machine vision

Image sequencing

Video consists of sequences of still images taken at high speed. Depending on the speed of the camera movement, one object is captured in many image frames. Therefore there is an redundancy of the information to be processed. First step in video processing is to find the amount of scene overlap in two subsequent frames.

Sanchiz et al. (1995) developed a system to sequence the image frames with assumption that there is no movement in the scene itself. However, in our case the plant

leaves were fluttering rapidly with wind and inclusion of plant region for overlapping frame gave erroneous result. The overlapped images were analyzed to find the plant count, inter-plant spacing, and plant centers. Tang and Tian (2002 a) developed an improved method of image sequencing similar to the method used by Shrestha and Steward (2001). They have implemented the sequencing process with optimizing the search area and using a look up table. This method could be used to implement the mosaicing process in real time up to vehicle speed of 1.21 m/s (4.4 km/h).

Segmentation and plant center location

The first step in machine vision for population sensing is the segmentation of the plant from the background. Image segmentation is the task of sub-dividing an image into constituent parts or objects (Gonzalez and Woods, 1992). After segmentation, each image pixel is assigned to one of several specific classes. In agricultural machine vision applications, segmentation is essential to separate plant from the background, i.e., soil and residue. Two general classes of methods have been used to separate a plant from its background in color images of agricultural crop fields

One class of methods transforms the pixels data into a one-dimensional index. Segmentation is then accomplished by thresholding this index's histogram. Meyer et al. (1998) segmented the plant and background by thresholding the excess green color index. Andreasen et al. (1997) segmented images by thresholding the median filtered histogram of the green chromaticity coordinates. Pérez et al. (2000) used a normalized difference index (NDI) along with morphological operations for plant segmentation. The motivation behind the NDI is that it is similar to a vegetative index commonly used in agricultural remote sensing to estimate the amount of vegetation represented by a pixel (Rees, 1999). NDI is given by the equation:

$$NDI = \frac{G - R}{G + R} \tag{1}$$

where G is a green channel value and R is red channel value. In order to determine the threshold value automatically, Otsu (1979) developed a statistical based optimum thresholding technique. The method maximizes the ratio of "between class variance" to "within class variance" of the normalized image. The drawback of this method is its rigidity

in threshold selection. The threshold cannot be adapted with changing lighting condition encountered in the field.

In field conditions under daylight, variability in lighting occurs (Steward et al., 1999), and hence practical segmentation algorithms must have the ability to adjust to such changes. Thus, a second class of methods treats the segmentation problem as a pattern recognition problem with the RGB values individually treated as class features. A Bayes classifier is then trained to accomplish segmentation by dividing up the color space with a decision surface. The use of such a classifier allows the training to be accomplished for individual images that represent various lighting conditions typically encountered in outdoor conditions. In addition, because of its general quadratic form, the decision surface produced by the Bayes classifier can take on many different shapes based on class statistics.

Tian and Slaughter (1998) used a Bayes classifier to do plant and weed segmentation with robustness to lighting variations. To train the classifier, individual pixels were first classified in a partially supervised fashion through cluster analysis. Then a Bayes classifier was trained so that a decision surface was defined to segment images with lighting conditions which are similar to those represented by the training image. Further refinements of this approach were documented by Steward and Tian (1998) who analyzed several different classification schemes to divide up the color space.

Besides using color for plant detection, different shape features of the plant can also be used to detect corn plants. Jia et al. (1990) studied the feasibility of detecting main veins along the leaves and found the intersecting point to estimate the corn plant center. However, at early growth stage of corn plant, there are no consistent distinct veins. Therefore, it was not possible to use main veins for plant center detection.

Nishiwaki et al. (2001) developed a template based plant center location algorithm for rice plants. They have used a Gaussian model for template formation and calculated the cross correlation of the template with near infrared (NIR) image of the rice field. This method however assumed that the plants are in uniform distance and did not take any account of doubles or missing plants.

Tang and Tian (2002b) used the 8-connected skeleton of the segmented binary image. With pre set conditions and using unsaturated color of the whorl, they were able to detect the corn plant center location up to V3 growth stage.

Stereo vision

Stereoscopic (stereo) vision has potential for vehicle-based, in-situ measurement of crop growth across crop fields. Stereovision is a machine vision technique for recovery of three-dimensional information using two imaging viewpoints. These viewpoints may come from two cameras spaced at some fixed baseline distance or from a single camera moved from one viewpoint to another between image acquisitions. Image formation in a camera is a many-to-one mapping from 3-dimentional real world coordinates to 2-dimentioanl image coordinates and thus depth in a single image is ambiguous. With stereo vision, depth is recovered through triangulation by using imaging geometry and the disparity between common scene points in the two images (Fig. 1). Stereo vision algorithms thus typically consist of three steps: a) camera calibration relating image pixels to a collection of points in the 3-D scene, b) point correspondence determining the pair of points in the two images which relate to a single point in the scene, and c) 3-D reconstruction – finding the 3-D location of points in the scene (Sonka et al., 1999).

Some work has been done on studying the feasibility of using machine vision to determine the plant height to show the relationship between plant height and other physiological relationship of a plant. Ling et al. (1996) found that the machine vision system was capable of determining a plant physiological response to the nutrient stress within 24 hours of the change of the nutrient regime. They have also found that the most rapid increase in top projected leaf area occurred from 4 to 5 hours after the onset of the dark period. Stereo vision has been applied into several agricultural applications. Tebourbi et al. (1999) developed a stereo vision system for measurement of soil texture and recommended that stereo vision be applied to crop growth sensing. McDonald et al. (1999) estimated the soil surface roughness and hence light reflection using stereo vision for different soil structures. They found good agreement between results obtained with the stereo vision approach and those obtained with other standard methods that were extremely time-consuming. Stereo vision has been applied to robotic harvesting applications to estimate the distance between

the fruit or vegetable and the robotic manipulator (Takahashi et al., 1998; Kondo et al., 1996). Stereo vision has also been used in light interception studies of several crops (Ivanov et al.,1994; Sinoquet et al., 1998).

Kanuma et al. (1998) showed that a cabbage growth analysis system which employed stereo vision improved measurement accuracy of leaf area over a single vision system. Matsuura et al. (2001) developed a transplant population growth analysis system that estimated average height, leaf area, projected leaf area, and mass volume with good correlation to destructive measurements. Lines et al. (2001) developed a stereo vision system which estimated the mass of free-swimming fish with a mean measurement error of 18%.

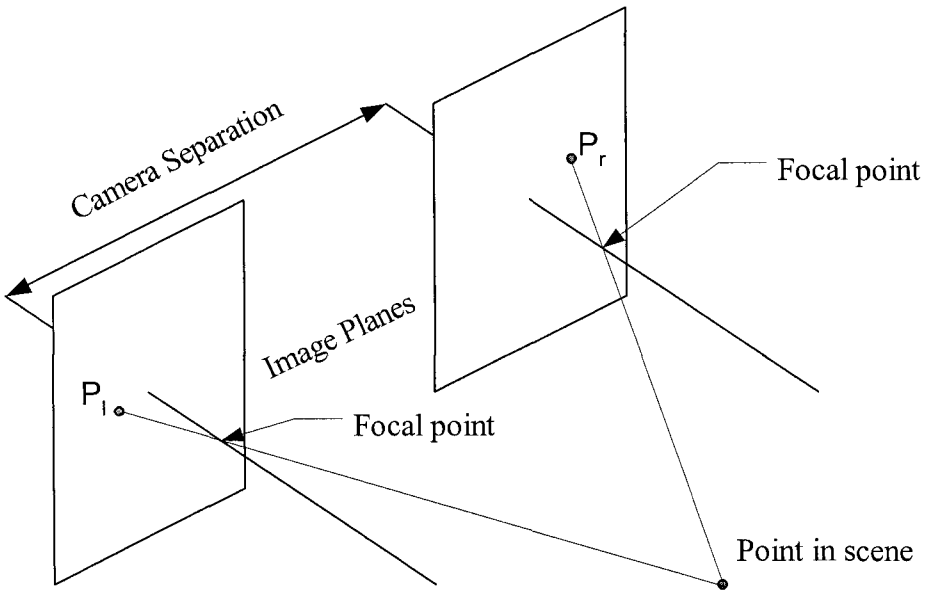


Figure 1. Illustration of the stereo vision concept. Depth is recovered from the scene through triangulation by using the disparity, $P_r - P_l$, of the projections of the scene point on the two image planes.

Will (2001) demonstrated that stereo imaging with vehicle-mounted sensors could estimate the 3-D world coordinates of control points with decimeter-level accuracy with a small number of images. The control points in Will's study were unambiguous and were determined manually in the images. However, Will demonstrated the potential of vision sensors and stereo-vision techniques on agricultural vehicles for estimating the location of physical points in field scenes.

Chapter 3. Segmentation of plant from background Using Neural Network Approach

A paper published in the Proceedings of Intelligent Engineering System through Artificial Neural Networks, 2001, Vol. 11, 903-908.
Dev S Shrestha[1], Brian L Steward[2], Eric Bartlett[3]

Abstract

This paper presents an artificial neural network approach to plant-background segmentation of agricultural field images. A truncated ellipsoidal (TE) surface was found to be most effective in defining the green region in RGB space. However, parameters defining the boundaries were subjective to the lighting condition and non-linear. A feed forward neural network was trained for segmenting early stage corn plant from background in various lighting conditions. Segmentation performance of the neural network approach was compared with normalized difference index (NDI) and Bayes classifier approaches. The neural network effectively estimated the parameters of the TE decision surface from first order image statistics and provided better segmentation performance than the other methods.

Introduction

Image segmentation is the task of sub-dividing an image into constituent parts or objects. (Gonzalez and Woods, 1993). After segmentation, each image pixel is assigned to one of several specific classes. Segmentation is a basic preprocessing task in many image processing applications. In agricultural machine vision applications, segmentation is essential to separate plant from the background, i.e., soil and residue. Two general classes of methods have been used to separate a plant from its background in color images of agricultural crop fields

One class of methods transforms the pixels data into a one-dimensional index. Segmentation is then accomplished by thresholding this index's histogram. Meyer et al. (1998) segmented the plant and background by thresholding the excess green color index. Andreasen et al. (1997) segmented images by thresholding the median filtered histogram of

[1] Graduate student
[2] Assistant Professor
[3] Associate Professor

the green chromaticity coordinates. Pérez et al. (2000) used a normalized difference index (NDI) along with morphological operations for plant segmentation. The motivation behind the NDI is that it is similar to a vegetative index commonly used in agricultural remote sensing to estimate the amount of vegetation represented by a pixel (Rees, 1999). NDI is given by the equation:

$$NDI = \frac{G - R}{G + R} \quad (1)$$

where G is green channel value and R is red channel value. In field conditions under daylight, variability in lighting occurs (Steward et al., 1999), and hence practical segmentation algorithms must have the ability to adjust to such changes. Thus, a second class of methods treats the segmentation problem as a pattern recognition problem with the RGB values individually treated as class features. A Bayes classifier is then trained to accomplish segmentation by dividing up the color space with a decision surface. The use of such a classifier allows the training to be accomplished for individual images that represent various lighting conditions typically encountered in outdoor conditions. In addition, because of its general quadratic form, the decision surface produced by the Bayes classifier can take on many different shapes based on class statistics.

Tian and Slaughter (1998) used a Bayes classifier to do plant and weed segmentation with robustness to lighting variations. In order to train the classifier, individual pixels were first classified in a partially-supervised fashion through cluster analysis. Then a Bayes classifier was trained so that a decision surface was defined to segment images with lighting conditions which are similar to those represented by the training image. Further refinements of this approach were documented by Steward and Tian (1998) who analyzed several different classification schemes to divide up the color space. Previously, algorithms were developed to specify Bayes classifier decision surfaces given field images taken under various lighting conditions. The next step in this research is to develop segmentation algorithms that can easily adapt to lighting conditions based on image-derived information.

Although the index approach is not typically thought of in terms of a decision surface, it does, however, functionally define a decision surface (Fig. 1). The only flexibility in the index approach is where the decision surface is positioned. While the classifier approach allows more flexibility in defining the decision surface, it does, however, require initial

labelling of pixel classes and estimation of the class statistics. These tasks are both computationally demanding and thus not well suited for real-time adaptation. A neural network approach of segmenting plant from background is presented in this paper. This approach has advantages over other approaches because it offers a flexible decision surface that can be adapted for lighting changes.

Objective

The objective of this research was to develop a segmentation method that could easily adapt to changes in lighting conditions. To validate this method, its segmentation performance was compared with that of a NDI index approach and a Bayes classifier approach.

Neural network method

This method accomplished segmentation by using an truncated ellipsoidal (TE) surface in RGB color space as a discrimination boundary between vegetation and non-vegetation regions. This surface was originally developed by plotting curves on constant B value planes to separate regions perceived as green from those perceived as non-green. The motivation behind this method was to roughly separate green and non-green regions in RGB color space and later fine tune this separation using a neural network. After observing this family of curves, it was determined that the decision surface could be functionally represented by a truncated ellipsoidal (TE) surface given by:

$$\frac{R^2}{D^2} + \frac{(1-G)^2}{(E \times B + F)^2} = 1 \qquad (2)$$

where R, G, and B values were the red, green and blue values of a particular pixel and D, E, and F were the parameters describing the shape of the ellipsoid. For a given set of parameters, the left-hand side of Eq. (2) was used to classify pixels as vegetation if ≤ 1 or background if >1. A TE decision surface for typical values of D, E and F is shown in Fig. 2.

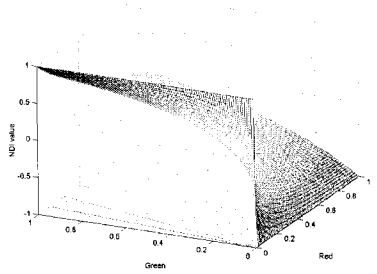


Figure 1. NDI index decision surface for different red and green values with contour lines of NDI values.

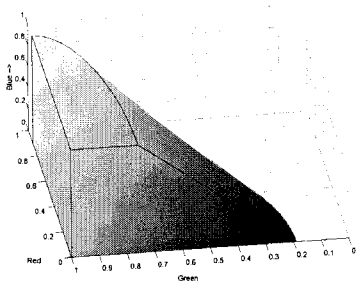


Figure 2. TE decision surface values with parameters D = 0.9, E = -0.57 and F = 0.81.

Pixel values are affected by the specific configuration of a color vision system, including factors related to intensity and spectral distribution of illumination, the lens and lens aperture, properties of RGB color filters, the image sensor response and the digitizer (Chang and Reid, 1996). Thus, the optimal values of the decision surface parameters will vary from image to image as lighting condition changes. Lighting will affect relative RGB values and their distribution. Hence, the first order statistics of the image RGB values were used to determine the optimal parameter values.

A fully connected 3-layer feed-forward neural network was developed to estimate the TE surface parameters as a function of the mean and standard deviation of an image's RGB values. Hyperbolic tangent sigmoidal activation functions were used in 2 hidden layers with 5 and 4 nodes respectively, and linear activation functions were used in the output layer. Resilient back propagation was used to find the minimum of error surface (Demuth and Beale, 1998). Initial weights of the neural network were assigned randomly.

Segmentation performance analysis

A set of 20 images was manually segmented by using visual inspection and the MATLAB AOI selection feature. These images were used for comparing the segmentation performance of the different methods. Segmentation performance was measured using segmentation error, which was defined as the number of pixels segmented into a different class then in the manually segmented image.

The goal of the first set of analyses was to ascertain if the decision surface associated with the three methods produced different performance when optimized for the image being segmented. For the NDI index segmentation approach, the individual NDI threshold for each of the 20 images which minimized segmentation error was found. To evaluate the performance of the Bayes classifier, the multivariate pixel statistics were estimated from each of the 20 manually segmented images. Then the estimated statistics for each image defined the classifier used to segment that image. For the TE decision surface, parameters values were found which minimized the segmentation error of each image. Average segmentation error for the optimized decision surfaces was compared to determine if statistical differences existed between them.

The second analyses evaluated segmentation performance of the three methods using constant parameter values for the entire set of twenty images. This was done to evaluate segmentation performance when a static segmentation scheme is used across images acquired under different lighting conditions. Five NDI threholds were selected within the optimum range of threshold parameters from 0.05 to 0.25 at 0.05 intervals to segment the 20 images. The average segmentation error for each threshold value was calculated.

For the Bayes classifier approach, a classifier trained from a randomly selected image was used to segment all 20 images, and the average segmentation error was calculated. This process was repeated 5 times. Similarly in order to determine parameter value sensitivity using the TE decision surface, the set of values optimized for a randomly selected image was used to segment all of the 20 images. This process was also repeated 5 times. Average segmentation error for each of the static methods was compared.

For the final analysis, optimum TE parameters were estimated manually for 190 images by a computer-aided method that gave the visually best segmentation of plant. 70%

of the images were used to train the neural network described above. The trained NN was applied to the remaining 30% of the 190 images as a validation set. The training and validation process was repeated 10 times by randomly selecting the images with the 7:3 ratio for training and validation. The NN output was compared with manually optimized values. A trained NN was also used to estimate the TE parameters for the 20 manually segmented images. Estimated TE parameters were then used modify the TE surface and then segment that image. Segmentation error was calculated for each segmented image.

Results and discussion

Statistically, the TE surface produced the minimum segmentation error of the three methods when each was optimized for individual images. The error for each method was significantly different than the others. Average segmentation error using NDI, Bayes and TE optimized decision surfaces were 6.43, 1.78 and 1.59 % respectively. These results were consistent with qualitative assessment of example image segmentations (Fig. 3). The optimum NDI threshold varied ranged from 0.05 to 0.25 across the 20 images.

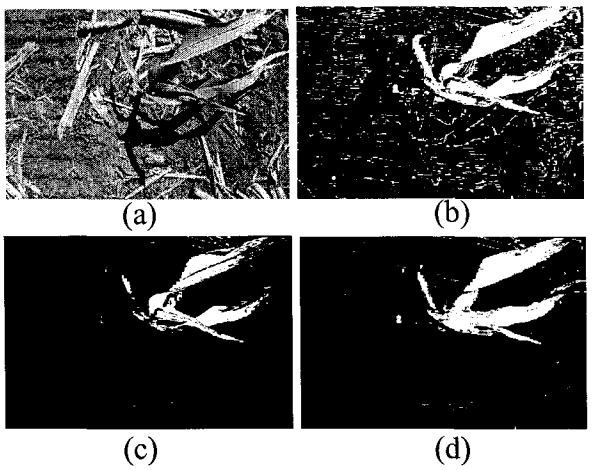


Figure 3. (a) Original image of a corn plant acquired with Sony DCR-TRV900 3CCD Mini DV camcorder. (b) Segmented Image using NDI method. (c) Segmented image using Bayes classifier. (d) Segmented image using neural network method. Optimized decision surfaces were used for these segmentations.

When a constant NDI threshold value was used for segmentation, average segmentation errors using threshold values from 0.05 to 0.25 were 12.3, 7.2, 7.7, 8.0 and 8.0% respectively. At the lowest threshold value, the number of background pixels classified

as plant pixels increased. Since there were more background pixels than plant pixels, the average error percentage was higher at the lower threshold value. The 0.10 threshold value resulted in a significantly lower segmentation error than the other threshold values. Segmentation using individually optimized threshold values resulted in significantly lower segmentation error (6.43%) than the best case static threshold value, 0.10.

Similar results were found for the Bayesian and TE methods. The minimum segmentation error obtained across the 5 trials using a static decision surface was 13.9% for the Bayes method and 3% for the TE method. These results confirm that dynamic decision surface parameters are required in order to optimally segment the vegetation irrespective of the method employed.

NN-estimated TE surface parameters were not significantly different than the manually estimated values. When the NN was used to dynamically adjust the TE decision surface parameters, the average segmentation error was 2.07% compared to 1.59% with manually optimized parameters. However, average segmentation error with dynamically adjusted parameters was significantly less than that associated with static decision surfaces.

Conclusions

The TE decision surface provided the best segmentation performance when each decision surface was optimized for individual images. When a static decision surfaces were to segment a whole set of images, segmentation error increased significantly for all methods. This result provided evidence that decision surfaces need to be adjusted for lighting conditions to achieve optimal segmentation.

An neural network which dynamically estimated TE decision surface parameters for individual images resulted in an average segmentation error which was significantly lower than the error obtained using a static parameters. Thus the neural network approach shows promise as a means of adaptively segmenting outdoor agricultural field images.

References

Andreasen, C., M. Rudemo, and S. Sevestre, 1997, "Assessment of weed density at an early stage by use of image processing," Weed Research, Vol. 37, pp. 5-18.

Chang Y.C. and J.F. Reid, 1996, "Characterization of a color vision system," Transactions of the ASAE, Vol 39, pp. 263-273.

Demuth, H. and M. Beale, 1998, Neural Network Toolbox for Use with MATLAB User's Guide Version 3, The MathWorks, Natick, MA, pp. 5.18 –5.20.

Gonzalez, R.C. and R.E. Woods, 1992, Digital Image Processing, Addison-Wesley Publishing Company, Reading, MA.

Meyer, G. E., T. Mehta, M. F. Kocher, D. A. Mortensen, and A. Samal, 1998, "Textural imaging and discriminant analysis for distinguishing weeds for spot spraying," Transactions of ASAE, Vol. 41, pp. 1189-1197.

Pérez A.J., F. López, J.V. Benlloch, and S. Christensen, 2000, "Colour and shape analysis techniques for weed detection in cereal fields," Computers and Electronics in Agriculture, Vol 25, pp. 197-212.

Rees, G., 1999, The Remote Sensing Data Book, Cambridge University Press, Cambridge, UK, pp. 245-6.

Steward, B. L. and L. F. Tian, 1998, "Real-time weed detection in outdoor field conditions," Proceedings, Precision Agriculture and Biological Quality, G. E. Meyer and J. A DeShazer, eds., SPIE., Bellingham, WA, pp. 266-278.

Steward, B. L., L. F. Tian, and L. Tang, 1999, "Detection of outdoor lighting variability for machine vision-based precision agriculture," ASAE Paper No. 99-3032, ASAE, St. Joseph, MI.

Tian, L. F. and D. C. Slaughter, 1998, "Environmentally adaptive segmentation algorithm for outdoor image segmentation," Computers and Electronics in Agriculture, Vol. 21, pp. 153-168.

Chapter 4. Automatic Corn Plant Population Measurement Using Machine Vision

A paper published in the Transactions of the ASAE, 2003, 46(2):559-565
D. S. Shrestha[1], B. L. Steward[2]

Abstract

A machine vision-based corn plant population sensing system was developed to measure early growth stage corn population. Video was acquired from a vehicle-mounted digital video camera at V3 to V4 stages under different daylight conditions. Algorithms were developed to sequence video frames and to segment, singulate, and count corn plants. Vegetation segmentation was accomplished using a truncated ellipsoidal decision surface. Two features were extracted from each pixel row of the segmented images: total number of plant pixels and their median position. Adjacent rows of the same class were grouped together and iteratively refined for final plant counting. Performance of this system was evaluated by comparing its estimation of plant counts with manual stand counts in 60 experimental units of 6.1 m sections of corn rows. The number of corn plants in these experimental units ranged from 14 to 48, corresponding to a population of 30,000 to103,000 plants /ha. In low weed field conditions, the system plant count was well correlated to manual stand count ($R^2 = 0.90$). Standard error of population estimate was 1.8 plants over 33.2 mean manual plant count or 5.4% coefficient of variation.

Keywords. *Image processing, precision agriculture, plant population, video cameras, corn, machine vision.*

Introduction

Corn plant populations that are higher or lower than optimal can reduce crop yield. Duncan (1958) found that corn yield was maximized at particular plant populations depending upon nutrient availability. Wiley and Heath (1969) investigated the relationships established by different researchers between corn population density and yield and found that the predictions had similar trends of yield maximization at particular plant population densities. Duncan (1984) presented the theory of crowding as a reason for yield reduction.

---

[1] Graduate Student
[2] Assistant Professor

However, optimum plant densities have not been constant over time, but have increased substantially over the last several decades (Troyer and Rosenbrook, 1983; Nafziger, 1994).

Even if a corn variety is planted at its optimal population, row spacing and interplant distance within a row can also affect the final yield. Plant population density, as well as interplant distribution, is important in effective utilization of available resources like nutrients and sunlight. Pablo et al. (2000) studied the row spacing effect at different levels of nitrogen availability in corn. They found that the corn yield was higher when the row spacing was decreased for the same population density. The relative yield increase was higher for nitrogen deficient fields. Doerge et al. (2002) measured spacing of 6,000 plants in research conducted in Missouri, Iowa and Minnesota. Whole field plant spacing standard deviation ranged from 3.2 to 6.9 inches. They estimated that every inch reduction in plant spacing standard deviation in a commercial field would increase the yield by about 3.4 bu/acre. Nafziger (1996) found that when there is a missing plant, the plants on either side compensated for only 47% of the reduced yield in lower population density fields (18,000 plants /acre) and 19% in higher plant density (30,000 plants/acre) fields, hence decreasing the final yield.

There are three main causes of variability in plant spacing: seed germination, planter seed placement, and plant death. Seed germination rates typically range from 90 to 95% (Nielsen, 2001). Planter performance depends both on maintenance of planter and speed. Nielsen (1995) reported that when the planter speed varied from 6.4 to 11.2 km/h (4-7 mph), the planted seed rate at higher speeds was significantly different than the planted seed rate at lower speeds. He concluded that a yield loss of at least 1.9 bu/acre occurs at every 1 mph speed increase in the range of 4-7 mph. Weather and pest related damage may result in unevenly spaced plant survivors within a row (Nielsen, 2001). Because of these factors, established plant population and spacing may be different than target plant population.

Bullock et al. (1998) found that for variable rate seeding to be profitable, a farmer needs extensive knowledge of site-specific plant population versus yield data from many years. Manual stand counts would not be feasible for a large field and is also susceptible to human error. An automated plant counting system provides a method for counting plants quickly and objectively. In addition, comparison of early stage plant population

measurements with populations at harvest can be used to measure the plant survival rate throughout the growing period. Plant survival rates could be used to estimate the population density required at planting time to achieve the desired population density at harvesting time.

Because of the importance of plant population density and spacing distribution to produce the optimum yield, several researchers have investigated population measurement systems. The majority of population sensing technologies have been developed for application at harvest. Birrell and Sudduth (1995) and Sudduth et al. (2000) have developed a combine mounted mechanical sensor to map corn population at harvest. Plattner and Hummel (1996) developed another corn population sensor using non-contact optical sensors at harvest. Nichols (2000) has also developed a corn population sensor using a moisture sensor to count corn stalks as they are pulled into the combine head. Easton and Easton (1996) developed a mechanical sensing system for counting young corn plants, which was mounted on a one-wheeled, human-powered cart.

With current advances in digital video technology, machine vision has potential as a sensing technology for corn plant population measurements. In addition, a machine vision system could also be extended to measure other field variables like plant color, soil color, plant height and other crop characteristics. Therefore machine vision was investigated as a means to sense corn plant population.

Automated plant counting using machine vision involves three major steps. First, the individual video frames must be separated and the amount of overlap of the scene in two subsequent frames must be determined in order to avoid multiple counting of plants that occur in two frames. Second, the plants must be segmented from the scene background. Third, plants must be singulated and counted.

The objective of this research was to develop a machine vision sensing system for counting corn population at an early growth stage ranging from V3 to V4. Specific research objectives were to:

1. Develop an image correspondence methodology for video frame sequencing that could reliably find the amount of shift from one frame image to the next in a video of field scenes.

2. Develop a corn plant segmentation and singulation algorithm that would accurately estimate corn plant population over row lengths.

Methodology

Experimental Setting

Two weeks after plant emergence, video sequences were collected in corn plots planted on April 26, 2001(Asgrow RX686RR) at the Iowa State University Agronomy and Agricultural Engineering Research Center, Boone, Iowa. A Sony DCR-TRV900 digital camcorder was mounted on a John Deere Gator utility vehicle at 0.60 m above the ground with a 0.30 m by 0.40 m field of view. Each captured image size was 480 × 720 pixels with 24 bit color resolution. The vehicle was driven over a corn row in a straight line with the camera directly over the plants at the speed of about 1 m/s. The shutter speed was adjusted to 1/1000 second, frames were captured in progressive scan mode, and other camera settings were set to auto. In the field, the video stream was recorded on a miniDV tape. The corn plants were at V3 to V4 growth stages, which are the vegetative growth stages of corn when the third or fourth leaf collar is visible. The corn row spacing was 0.76 m (30 in.), and the target population was 74,000 plants/ha (30,000 plants per acre). Corn rows were divided into 6.1 m (20 ft.) long sections by staking yellow construction tape perpendicular to the row direction. Each 6.1 m long corn row section was considered to be an experimental unit, and a total of 60 experimental units were used in the study. This length represented a trade-off between population measurement resolution and spatial sampling resolution. In addition, 6.1 m row sections are slightly longer than the row length recommended to achieve the recommended 1/1000-acre stand counts for 0.76 m (30 in.) rows (Benson, 1990). The number of plants within each experimental unit was determined through manual stand counts.

In the laboratory, video streams were transmitted from the camera to a personal computer using an IEEE 1394 serial interface. Adobe Premiere® 6 (San Jose, CA) was used to capture the video stream as AVI files and then to decompress and store individual frames as color tagged image file format (TIFF) files. Matlab® Ver. 6 (The Mathworks, Inc., Natick,

MA) was used for development of image processing algorithms and subsequent image processing.

Image Sequencing

Image sequencing is the process of determining the amount of overlap in succeeding video frames. This is essentially an image correspondence problem where common scene points in two images are identified and matched. There are many methods available in the literature for image correspondence. One technique is to match a pattern and a searched image through the use of a matching criterion that serves as a measure of correlation (Sonka et al., 1998). Feature-based image correspondence such as the method developed by Dai and Khorram (1999) is another possible approach for matching remotely sensed image pairs. Their algorithm included image segmentation, control-point selection and correspondence, and transformation parameter estimation. In general, for the feature-based algorithms to be effective, images should contain objects with well-defined shapes and edges like a river or road as usually encountered in case of remote sensing. Feature-based algorithms also tend to be computationally expensive. For cornfield scenes, the objects are not well defined. In addition, the computation time must be constrained due to the large number of frames to be sequenced. Sanchiz et al. (1995) developed a feature-based system to sequence the video frames in fields containing small cabbage plants with the assumption that there is no movement in the scene itself. In our case, however, the corn plant leaves were moving in the wind, and inclusion of plant regions in the correspondence algorithm produced erroneous results.

Image correspondence can be done both in spatial and frequency domains. In the frequency domain, image correspondence can be obtained to sub-pixel accuracy, but the computational cost is higher than spatial correlation-based image matching (Averbuch and Keller, 2002). Correspondence is a key problem in machine vision applications and no general reliable solution exists (Maciel and Costeira, 2002).

In this research, to accomplish image sequencing, intensity images were derived from color images, and the amount of shift between sequential frames was estimated. Assuming that camera rotation was negligible from frame-to-frame, the image sequencing problem consisted of finding the shift in the next frame relative to the current frame being processed.

A patch was selected randomly in the current frame with the constraint that its expected corresponding matching location (search region) was within the boundary of the next frame. A 30×30 pixel patch was selected as a balance between two competing criteria: to minimize computation time and to maximize textural content which both increase with patch size. This patch size corresponded to 1.88 cm in the direction of travel and 1.67 cm perpendicular to the direction of travel. When selecting a patch in the current frame for matching with next frame, the anticipated shift was taken into account so that the search region in the next frame completely lay within the next frame boundary (Fig. 1). The search region was set such that the patch could be moved by 30 pixels in any direction from the anticipated amount of shift. The average shift of two previous images was used to determine the anticipated shift amount. Once a patch in the current frame and corresponding search region were selected, both patch and search region were searched for vegetation by segmenting plants using the truncated ellipsoid method, described in the next section. If more than 5 % of the pixels in either patch or search region were classified as vegetation, than that patch selection was disqualified from further processing. In addition, if the patch was very dark (average intensity < 0.2) or very bright (average intensity > 0.8), the patch was disqualified and another patch was reselected randomly. Plant regions were excluded because the position of a plant may change from frame to frame due to wind moving the leaves leading to a false match. If the intensity of a patch was too high, generally it was too saturated to contain texture information. Similarly, dark patches usually had a low level of information.

For the first two images in a sequence, there was no information available for the anticipated shift. Therefore, to determine the amount of shift between the first and the second frame in the sequence, it was assumed that the vehicle always traveled forward and the patch was selected within the lower 100 rows of the first frame within a 50-column margin from both right and left sides. Then the entire second frame was searched for the match location of that patch.

If the patch was m×n pixels and search region was M×N pixels, the matching error for each position was determined by:

$$Err_{p,q} = \sum_{i=1}^{n} \sum_{j=1}^{m} \left| P_{i,j} - S_{i+p-1,j+q-1} \right| \quad (1)$$

where, Err is the M-m × N-n error matrix. The (p, q) term of Err corresponds to the sum of absolute errors when the patch was shifted by (p, q) pixels from the upper left corner of search region. P is the intensity patch from the current frame, and S is the search region from the next frame (Fig. 2).

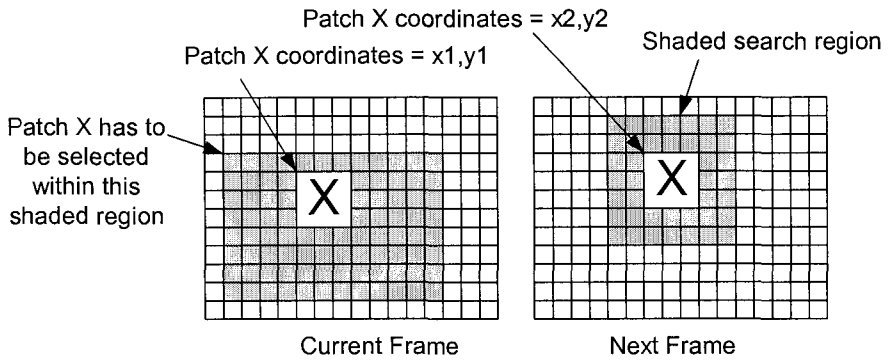


Figure 1. For image sequencing, an image patch X in the current frame was shifted in a search region in the next frame in search for the best match. The difference in coordinates of the patch matched to the second frame gives the amount of shift from the current frame to the next.

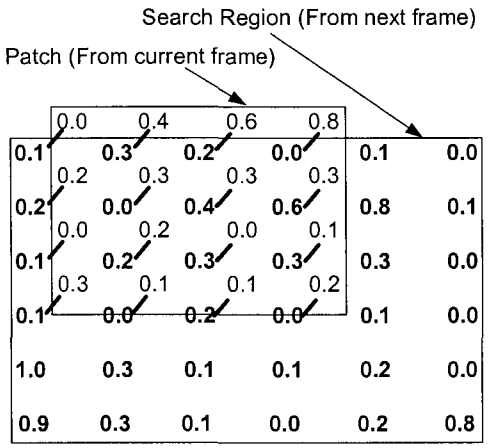


Figure 2. Process of calculating an error matrix, Err. The patch was shift over the search region. For the position shown above $Err_{1,1}$ = |(0.0-0.1)|+|(0.4-0.3)|+...+|(0.2-0.0)| = 3.3.

A candidate match was found by finding the minimum valued element in Err. To determine the validity of a match, the minimum value of Err had to be significantly lower than other values (Fig. 3). In order to test for a statistically significant minimum, the values in Err were sorted in ascending order, and the difference between successive values was

calculated. For a valid match, the difference between the lowest error and the next to the lowest error value was required to be higher than 5 standard deviations (σ) from the mean of the remaining error differences. For example, the error matrix for figure 2 was calculated as:

$$\mathrm{Err} = \begin{vmatrix} 3.3 & 4.3 & 4.3 \\ 2.8 & 0.0 & 3.0 \\ 4.1 & 2.6 & 3.8 \end{vmatrix} \tag{2}$$

The matrix Err was rearranged in a row of ascending values, and the difference ΔErr was calculated as:

$$\Delta\mathrm{Err} = \begin{bmatrix} 2.6 & 0.2 & 0.2 & 0.3 & 0.5 & 0.3 & 0.2 & 0.0 \end{bmatrix} \tag{3}$$

Since the first value of ΔErr i.e. 2.6 is more than 5 standard deviations from the mean of the rest of the differences, the minimum error 0.0 in the Err matrix was considered to be a true minimum and the match was accepted. If a valid match, based on a 5 σ criteria, could not be found in the specific region, then another random patch was chosen in current frame and searching was repeated.

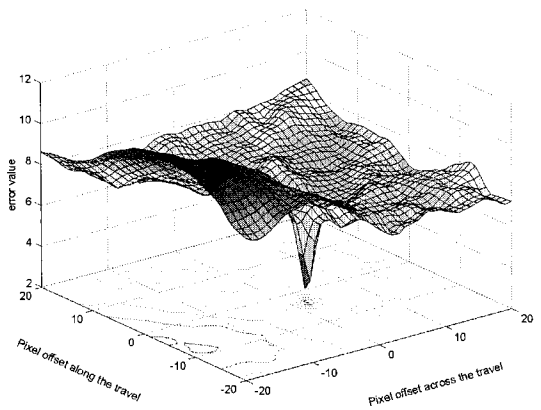


Figure 3. Example of an error surface and its contour for a typical patch matching. The minimum error value must be significantly different from the mean error to be accepted as a valid match.

This criteria was established from Chebyshev's theorem which applies to any distribution (Walpole and Myers, 1978) and shows that the minimum probability of a correct match is given by:

$$P \geq 1 - \frac{1}{5^2} = 0.96 \tag{4}$$

where P is probability that the value of the random variable will be less than 5 standard deviations from mean. Thus for error differences that are more than 5 standard

deviations from the mean, there is at most a 4 % probability that the match was due to a random minimum in the error matrix.

Image Segmentation

The next step after image correspondence was segmentation of vegetation from background. There are different methods available for separating vegetation from non-vegetation region. Meyer et al. (1998) segmented plant and background by thresholding the excess green color index. Andreasen et al. (1997) segmented images by thresholding the median filtered histogram of the green chromaticity coordinates. Perez et al. (2000) used a normalized difference index (NDI) along with morphological operations for plant segmentation. The segmentation algorithm employed in this research should be able to segment the plant in changing lighting conditions that occur in the field due to clouds and the time of the day.

Tian and Slaughter (1998) developed an algorithm to achieve segmentation robustness in outdoor field images under varying lighting conditions. This algorithm was based on cluster analysis of pixels in color space for labeling pixels and Bayesian classification for the development of a decision surface in color space. The segmentation method used for this research employed a decision surface in color space that was defined by only three parameters (Shrestha et al., 2001). This surface was a truncated ellipsoidal surface given by:

$$\frac{R^2}{D^2}+\frac{(1-G)^2}{(E\times B+F)^2}=1 \qquad (5)$$

where R, G, and B were the red, green, and blue intensities ranging from 0 to 1, and D, E, and F were the parameters describing the shape of the ellipsoid. Each of these parameters has a physical meaning based on the perceived green region in color space. D is the maximum red intensity still perceived green when B = 0 and G =1. E is the slope of the ellipsoid boundary in the green-blue plane. F is the distance from maximum to minimum green intensity that is perceived green when both blue and red channels are zero (Fig. 4). Constant parameter values D = 0.9, E = -0.57 and F = 0.81 were used for image segmentation in this research. These parameter values were determined by Shrestha et al. (2001) to provide

a general segmentation of plants across varying outdoor lighting conditions. For a given pixel color vector, the pixels were classified according to the decision rule:

$$\frac{R^2}{D^2}+\frac{(1-G)^2}{(E\times B+F)^2}\underset{\omega_2}{\overset{\omega_1}{\gtrless}} 1 \tag{6}$$

where ω_1 is the background class and ω_2 is the vegetation class. Because the surface is only defined by three parameters, adjustments can easily be made as lighting conditions change. Adjustments of the surface parameters by a neural net have been investigated and are the topic of another paper (Shrestha et al., 2001).

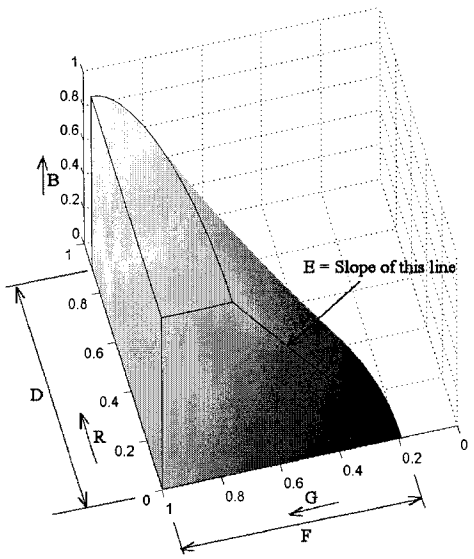


Figure 4. The truncated ellipsoidal decision surface in RGB color space used to segment vegetation from background. A pixel with RGB intensities inside the ellipsoid is considered a plant pixel. Parameter values used to segment plant from background were: D = 0.9, E = -0.57 and F = 0.81.

Plant Counting

After image sequencing and segmentation, images consisting of sequenced frames were analyzed to determine the number of plants and plant center locations. Jia and Krutz (1992) studied the feasibility of detecting main veins along leaves and found the intersecting point to estimate the corn plant center. At early growth stages, however, there were no consistent distinct veins observed in corn plant leaves. Therefore, it was not possible to use main veins for plant center detection.

In order to determine the plant centers and to count the plants, two features were extracted from every row of the binary segmented images: 1) the total number of plant pixels in each image row and 2) the median position of the plant pixels along each row. Once all the image rows were scanned and extracted features were recorded, a row was either classified as a plant row or a background row. An image row was classified as a plant row if:

1. the variation in median position of that row to the previous was less than the total number of plant pixels in that row and,
2. the plant pixel count of that row was greater than the mean value of total plant pixels in each row across the entire experimental unit.

Once a frame sequence from an experimental unit had been initially classified, adjacent plant rows and background rows were grouped into plant or background regions, and the average length of plant and background regions were calculated. Plant center row locations were estimated to be the middle row of each plant region. The plant center column was the mean of the median positions for that region. This classification resulted in an initial estimate of the number of plants and plant center locations. Next the plant and background regions were further refined using the following rule base.

1. Plant regions that were less than 20% of the mean plant region length were considered to be false plant regions and were reclassified as background.
2. Background regions that were less than 20% of the mean background region length were considered to be false background regions and were reclassified as plant.
3. Any plant center found outside a five standard deviations interval from the mean plant center position across the sequence was considered to be a weed and thus the region is reclassified as background.

After this refinement, the plants were counted again. If the plant count varied by more than 5% of the original count, the plant and background statistics were updated, and the regions were refined again through the rule base. When the plant count varied by less than 5%, the algorithm stopped.

Finally, plant regions that were more than twice the length of an average plant region were counted as doubles and more than three times as triples. The plant center row locations

were assumed to be at the middle of the plant region. In cases of multiple plants, the center was assumed to be at the middle of each of the adjacent plant regions.

Experimental Design

Image sequencing performance

To evaluate the performance of the image sequencing algorithm, a video sequence of 50 images was processed with the algorithm. Processing was repeated 30 times. For each image pair, the number of failed patch and search region selection attempts, number of attempts to achieve a significant match, and shifts along and across the direction of travel between subsequent frames were all recorded. The SAS (SAS Institute, Cary, N.C.) General Linear Model procedure (GLM) was used to test for significance differences in shifts statistics across replication of the algorithm and image pairs.

Plant counting performance

The plant count estimated by the sensing system was compared with that measured manually in 60 experimental units. Linear regression analysis was used to analyze the relationship between the two measurements. To determine the false plant and background region thresholds used in the algorithm, they were varied from 5 to 25% at 5 % intervals. For each threshold combination, the plant counting algorithm was used to estimate the number of plants in 13 randomly chosen sequences. The estimated plant count was compared to the manual count and a sum square of error statistic was calculated. The combination that minimized the error statistic was found and was used in the analysis of overall counting performance. The sensitivity of the algorithm's accuracy to the false plant and background thresholds as well as to the algorithm stopping criteria and distance from the crop row threshold were analyzed using all 60 experimental units. The plant count in each of 60 experimental units estimated by the sensing system was compared with that measured manually. Linear regression analysis was used to analyze the relationship between the two measurements.

Results and Discussion

Image sequencing performance

From the analysis of variance of the 50 frame sequence, the mean shift between two images along the direction of travel was 71.58 pixels, which, based on the camera field of view, would be a shift of 0.045 m of the field surface between two images. From this shift distance, the vehicle speed was estimated to be 1.34 m/s. At this speed, 85 percent of each frame is overlapped with the previous frame. There were significant differences in the mean pair-to-pair shifts in the travel direction ($P < 0.0001$). These differences were expected and are due primarily to variations in vehicle speed. After accounting for the pair-to-pair variation, the standard deviation in the shift estimation process was 2.3 pixels. This corresponded to 0.0014 m of the field surface or 3.2 % of the mean shift. The modified Levene test (Conover et al., 1981) revealed significant differences in shift variance across image pairs ($P < 0.0001$). This finding indicated that the image correspondence algorithm was finding larger differences across replications of the algorithm in particular image pairs than in other image pairs. Since the location of the patches is random, correspondence of frame pairs may vary depending on the location of the patch. This variation is due partly to uneven depth to objects in the image scene and vehicle yaw across the pair resulting in a rotation from one image to the next. The algorithm operated under the assumption of negligible frame-to-frame camera rotation, and based on these results, any error introduced because of this assumption was small.

The mean lateral shift between two images was 0.58 pixels, revealing that either the vehicle was turning or that the camera was slightly rotated relative to the centerline of the vehicle. After accounting for the pair-to-pair shift with the ANOVA model, standard deviation of the shift estimation algorithm was 1.7 pixels. Once again, there were significant differences ($P<0.0001$) in the shift variance across image pairs.

The number of attempts required to achieve a significant match ranged from 1 to 5 across the entire experiment. The mean was 1.21 attempts and standard deviation was 0.49 attempts. There were significant differences in the number of attempts ($P < 0.0001$) indicating that some image pairs tended to require more attempts than others. Upon further examination of these pairs, often one of the frame images was blurred leading to difficulty in

establishing a significant match. Nevertheless, even though on the average more matching attempts were required for blurred images than sharply focused image, significant matches were found for each of the pairs in the sequence. In addition, only 1.9 percent of the initial patch selections were rejected. A maximum number of two rejections before final patch selection occurred in 0.20% of the cases. Even when randomly selected patch was valid, the initial search region selection was rejected 5 % of the time, hence forcing the algorithm to reselect for a new patch. The search region area was nine times larger so an increased likelihood of finding a plant region in a search region was expected.

Plant counting performance

Manual plant counts over the 60 corn row sections varied over a range of 14 to 48 plants with a mean value of 33.2 plants. Linear regression analysis resulted in a R^2 value of 0.90 (Fig. 5). The slope and intercept of the regression line were 0.93 and 1.98, respectively, and were not significantly different from 1 and 0 respectively. The RMS error of plant counts estimated by the system was 1.8 plants over the 6.1 m length of a corn row. This error was 5.4 % of the mean. Based on the manual counts, the local population in the corn row sections varied from 30,100 to 103,000 plants/ha (12,200 to 41,800 plants/acre). Measuring over a 6.1-m corn row with 0.71m spacing resulted in a measurement resolution of ± 2150 plants/ha which was 3% of the target population.

From the analysis of the set of 13 sequences, the combination of a false plant region threshold of 20% of the mean plant region length and a false background region threshold of 20% of the mean background region length gave the least squared error. This combination was thus selected and used to count plants in all 60 experimental units in the analysis of overall system performance. System accuracy was sensitive to the false plant and background thresholds. However, an R2 greater than 0.8 was found when the false plant threshold was within the range of 10 to 20% and that for false background was between 15 to 25% (Fig. 6). A maximum R2 of 0.9 occurred with the 20%-20% threshold combination.

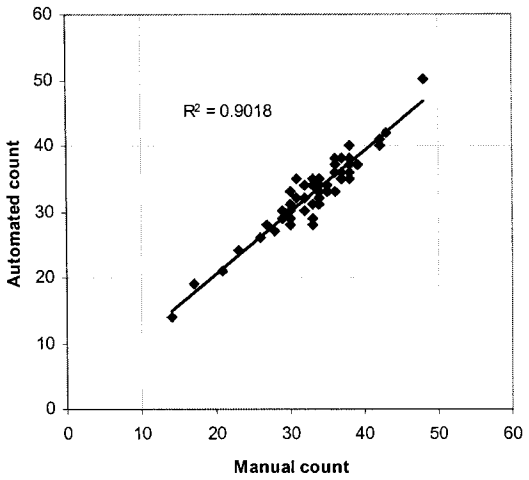


Figure 5. Regression of system estimated counts onto manual counts for 60 experimental units. The regression had an R^2 of 0.9018 and an RMSE of 1.8 plants.

Plant count accuracy was not sensitive to variations in the threshold for excluding plant pixels. When this threshold was varied from 3 to 5 standard deviations away from mean row position, R^2 varied by 0.902 to 0.896. The plant count algorithm-stopping criterion was also varied to investigate its effect on the number of refinement iterations and counting accuracy. When a change in plant counts less than 5% was required, it took an average of 3.0 iterations before the refinement algorithm stopped. When the stopping criterion was increased to 10%, the mean number of iterations decreased to 2.53, but R^2 also decreased to 0.84. When the stopping criterion was changed to 1%, the mean number of iterations increased to 3.38 times before stopping, but R^2 only increased to 0.906. Therefore, the stopping criterion of 5% was found to be a suitable tradeoff between accuracy and time.

One of the main sources of error found was variability in plant size and leaf orientation within an experimental unit. This made the threshold used to refine the plant and background region sensitive to plant size distribution. More weed and noise pixels were counted as plants when the false plant region threshold was lowered below 20% and small plants were considered to be weeds when the threshold was increased. However, under low weed conditions and plant growth stage of V3-V4, the system was able to estimate the number of plants across in a 6.1 m row with an RMSE less than 3 plants over a range of parameters.

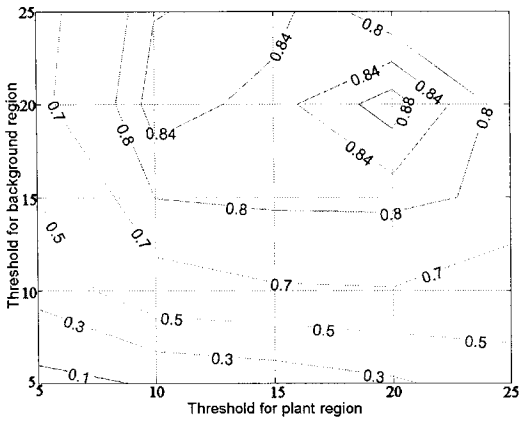


Figure 6. Contour plot of R^2 for different combinations of plant and background region thresholds. Thresholds between 10 and 20 for the plant region and thresholds between 15 and 25 for the background region produced $R^2 > 0.8$.

Conclusions

1. A patch matching algorithm with criteria for region selection and match validity is a feasible method for sequencing the video frames of corn row scenes acquired by a commercial digital video camera on a vehicle moving at 1 to 2 m/s.

2. A plant counting algorithm using two easily obtainable image features and a straightforward iterative rule-base was able to achieve population measurement accuracies similar to the system measurement resolution.

References

Andreasen, C., M. Rudemo, and S. Sevestre. 1997. Assessment of weed density at an early stage by use of image processing. *Weed Research* 37: 5-18.

Averbuch, A. and Y. Keller. 2002. FFT based image registration. In *International Conference on Acoustics, Speech, and Signal Processing* 4: 3608-3611.

Benson, G.O. 1990. Corn replant decisions: a review. *Journal of Production Agriculture* 3(2): 180-184.

Birrell, S.J. and K.A. Sudduth. 1995. Corn population sensor for precision farming. *ASAE Paper No. 951334.* St. Joseph, MI: ASAE.

Bullock, D.G., D.S. Bullock, E.D. Nafziger, T.A. Doerge, S.R. Paszkiewicz, P.R. Carter, and T.A. Peterson. 1998. Does variable rate seeding of corn pay? *Agronomy Journal* 90(6): 830-836.

Conover, W. J., M. E. Johnson, and M. M. Johnson. 1981. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 23(4): 351-361.

Dai, X. and S. Khorram. 1999. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Transactions on Geoscience and Remote Sensing* 37(5): 2351-2362.

Doerge, T., T. Hall, and D. Gardner. 2002. New research confirms benefits of improved plant spacing in corn. *Crop Insights* 12(2): 1-5.

Duncan, W.G. 1958. The relationship between corn populations and yield. *Agronomy Journal* 50: 82-84.

Duncan,W.G. 1984. A theory to explain the relationship between corn population and grain yield. *Crop Science* 24(6): 1141-1145.

Easton, D.T. and Easton, D. J. 1996. Device to measure and provide data for plant population and spacing variability. U.S. Patent No. 5,568,405.

Jia, J., and G.W. Krutz. 1992. Location of maize plant with machine vision. *Journal of Agricultural Engineering Research* 52(3): 169-181.

Maciel, J. and J. Costeira. 2002. Robust point correspondence by concave minimization. *Image and Vision Computing* 20(9-10): 683-690.

Meyer, G.E., T. Mehta, M.F. Kocher, D.A. Mortensen, and A. Samal. 1998. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Transactions of the ASAE* 41(4):1189-1197.

Nafziger, E.D. 1994. Corn planting date and plant population. *Journal of Production Agriculture* 7(1): 59-62.

Nafziger, E.D. 1996. Effects of missing and two-plant hills on corn grain yield. *Journal of Production Agriculture* 9(2): 238-240.

Nichols, S.W. 2000. Method and apparatus for counting crops sensor. U.S. Patent No. 6,073,427.

Nielsen, R.L. 1995. Planting speed effects on stand establishment and grain yield of corn. *Journal of Production Agriculture* 8(3): 391-393.

Nielsen, R.L. 2001. Stand establishment variability in corn. AGRY-91-1 (Rev. Nov.-01). *Department of Agronomy*, Purdue Univ., W. Lafayette, IN.

Pablo, A.B., S.R. Hernan, H.A. Fernando, and E.E. Hernan. 2000. Row spacing effect at different levels of nitrogen availability in maize. *Agronomy Journal 92: 283-288.*

Perez, A.J., F. Lopez, J.V. Benlloch, and S. Christensen. 2000. Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture* 25: 197-212.

Plattner, C.E. and J.W. Hummel. 1996. Corn plant population sensor for precision agriculture. In *Proceedings of the Third International Conference on Precision Agriculture.* eds. P.C. Robert, R.H. Rust, W.E. Larson. June 23-26, Bloomington, Minnesota USA. 785-794.

Sanchiz, J.M., F. Pla, J.A. Marchant and R.S. Brivot. 1995. Structure from motion techniques applied to crop field mapping. *Image and Vision Computing* 14: 353-363.

Sonka, M., V. Hlavac, and R. Boyle. 1998. Segmentation. In *Image processing, analysis and machine vision.* Ch. 5, 190-193. Pacific Grove: PWS Publishing.

Shrestha, D.S., B. Steward, and E. Bartlett. 2001. Segmentation of plant from background using neural network approach. In *Intelligent Engineering Systems Through Artificial Neural Networks, Proceedings of Artificial Neural Networks in Engineering (ANNIE) International Conference.* 11: 903-908. St. Louis, MO.

Sudduth, K.A., S.J. Birrell, M.J. Krumpelman. 2000. Field evaluation of a corn population. In *Proceedings of the Fifth International Conference on Precision Agriculture.*, eds. T.C. Robert, R.H. Rust, W.E. Larson. July 16-19, Bloomington, MN.

Tian, L.F. and D.C. Slaughter. 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture* 21(3): 153-168.

Troyer, A.F. and R.W. Rosenbrook. 1983. Utility of higher plant densities for corn performance testing. *Crop Science* 23: 863-867.

Walpole, R.E., R.H. Myers. 1978. Random variables. In *Probability and statistics for engineers and scientists (2nd edition).* Ch. 2, 72-75. New York: Macmillan Publishing Co.

Wiley, R.W., and S.B. Heath. 1969. The quantitative relationship between plant population and crop yield. *Advances in Agronomy* 21: 281-321.

Chapter 5. Image Processing Algorithms for Early Stage Maize Plant Detection

A paper submitted to the Biosystems Engineering Journal (formerly Journal of Agricultural Engineering Research)

D. S. Shrestha[1], B. L. Steward[2], S. J. Birrell[3]

Abstract

Algorithms were developed to process video of maize rows and extract plant features to estimate density and spacing of early growth stage maize plants. Otsu's method was modified to compensate for noise present in threshold calculations. Three features were extracted and used to differentiate between weeds and maize plants: projected plant canopy area, plant length in the image row direction, and perpendicular distance of estimated plant center from the mean crop row position. Algorithm performance was analyzed across three tillage treatments, three growth stages from V3 to V8, and three population densities varying from 27,000 to 81,500 plants/ha. Overall, the algorithm estimated the number of plants in 6.1 m crop row lengths with an RMSE of 2.1 plants or 8.7% of the mean manual count of 24.1 plants per experimental unit. The mean measurement error was significantly different across tillage treatments, but no evidence of significant differences was found across growth stages and plant populations. The error variance at V7-V8 growth stages was significantly higher than that at V3-V4 growth stages. No significant differences were found between mean measured and estimated plant spacing distances.

Keywords: Image processing, maize population, precision agriculture, Otsu's method, MLE.

Introduction

The actual interplant distance in crop rows may be different than targeted plant spacing. Even if actual average plant population over a large field matches with the targeted plant population, uneven plant spacing adversely affects the yield (Doerge et al., 2002). Plant population and plant spacing variability, have an important effect on maize yield (Nafziger,

---

[1] Graduate Student
[2] Assistant Professor
[3] Assistant Professor

1996). Doerge et al. (2002) reported that every 2.5 cm (1 in) reduction in maize plant spacing standard deviation resulted in yield increases of about 0.21 Mg/ha. Nafziger (1996) found that the maize plants on either side of a missing plant compensated for only 47% of the reduced yield in lower population density fields (44,500 plants /ha) and 19% in higher plant density (74,000 plants/ha) fields, hence decreasing crop yield.

There are three main causes of variability in plant spacing: seed germination, planter seed placement, and plant death (Neilsen, 2001). Seed germination rates typically range from 90% to 95% (Nielsen, 2001). Planter performance depends both on planter maintenance and speed. Nielsen (1995) reported that when the planter speed varied from 6.4 to 11.2 km/h, the planted seed rate at higher speeds was significantly different than the planted seed rate at lower speeds. He concluded that a yield loss of at least 86 kg/ha occurs at every 1 km/h speed increase in the range of 6.4 to 11.2 km/h. Weather and pest-related damage may result in unevenly spaced plant survivors within a row (Nielsen, 2001). Because of these factors, established plant population and spacing may be different than targets.

Bullock et al. (1998) reported that a farmer needs extensive knowledge of site-specific plant population and yield data for many years for variable rate seeding to be profitable. Manual stand counts are not feasible for a large field and are also susceptible to human error. In addition, except for high value crops, variable rate application (VRA) technology has not been as widely adopted as was originally anticipated by experts. Bullock et al. (2002) presented an economic model showing that VRA fertilizer application is not yet profitable because two elements are still lacking: (1) the understanding of the relationship between yield and managed and unmanaged field variables and (2) the absence of low cost, accurate field variable measurement technology. If plant population and spacing variability could be measured more extensively, the understanding of their effect on yield could be increased.

Birrell and Sudduth (1995) developed a combine mounted mechanical sensor to map maize population at harvest which was an excellent estimator of hand counted population (Sudduth et al., 2000). Plattner and Hummel (1996) developed another maize population sensor using non-contact optical sensors at harvest. Shrestha and Steward (2003) demonstrated that machine vision can be effective in locating maize plants and measuring

interplant spacing from videos of crop rows. In their work, a manually selected threshold was used to classify plant and background regions and no attempt was made to distinguish maize plants from weed plants.

The overall goal of the statistical approach described in this paper was to generalize these previous image processing algorithms to singulate maize plants in a wider range of field conditions with increased robustness. Statistical approaches were developed to distinguish between weed plants, and single, double and triple maize plants. The particular objectives of this paper were to 1) develop an algorithm, which can be used in typical field conditions, to identify, locate, and count early growth stage maize plants using feature statistics, and 2) validate the algorithm performance across varying tillage, growth stage, and population factors.

Methodology

Video of maize rows was collected across different field conditions. The algorithm processed video by sequencing video frames, extracting image row features, and classifying segmented vegetation into maize plants or weeds (Fig.1). Algorithm performance was analyzed by comparing its estimation results with manual measurements.

Algorithm development

Frame Sequencing

Frame sequencing is the process of determining the amount of spatial overlap in succeeding video frames. This image correspondence problem involves identification and matching of common scene points in two frames. Frame sequencing was necessary to discard duplicate information and prevent multiple counting of maize plants.

An area correspondence algorithm developed by Shrestha and Steward (2003) was used to sequence video frames by estimating spatial shifts from one frame to another in both the frame column and row directions (Fig. 2). When the frame columns were not parallel with the crop row direction, measurement of plant center location from the frame edge had to be adjusted to determine the position from the line parallel to the row direction. The angle between frame columns and the crop row vector was due to camera rotation caused by: a) camera misalignment with the centerline of the vehicle and b) vehicle yaw. Camera

misalignment introduced a constant frame-to-frame shift along image rows. Vehicle yaw introduced a variable frame-to-frame shift along image rows causing the plant center location along image rows to vary when measured from the left edge of the image. Vehicle yaw also introduced a rotation of the frames relative to the direction of the crop row (Fig. 2).

Camera rotation, θ, was estimated by taking the tangent of the ratio of total displacement in the x direction for each video segment , X_T, and total displacement in the y direction, Y_T (Fig. 2). The coordinates of each plant center in the frame coordinate system were transformed into a coordinate system whose coordinates were parallel and perpendicular to the crop row. This transformation was accomplished with a rotational matrix:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (1)$$

where x' and y' were plant center locations in the crop row coordinate system, and x and y were plant center location in the image coordinate system. All distance measurements were in pixels. Since interplant distance were manually measured along the crop row, this transformation was necessary to estimate actual interplant distances along the same direction.

Plant and Background Region Classification

After sequencing, the portion of each frame not overlapped with previous frames was segmented into vegetation and background regions using the truncated ellipsoidal method (Shrestha et al., 2001). Two row features were extracted from each segmented frame row: 1) the total number of vegetation pixels and 2) the median position of the vegetation pixels. An image row 'i' was classified as a vegetation row if 1) the absolute difference between the median positions of row 'i' and 2) median position of row 'i+1' was less than total vegetation pixel count for row 'i', and row 'i' had more vegetation pixels than the mean number of vegetation pixels per row of all frame rows of a video segment containing vegetation pixels. This preliminary classification was based on the assumption that the number of noise pixels in a frame row associated with a background region would be less than that in a vegetation region. This initial classification of rows was followed by grouping consecutive plant rows into vegetation regions and consecutive background rows into background regions. Two additional features were determined for each vegetation region: 1) the total number of

vegetation pixels or "canopy area" and 2) the median vegetation location in two directions (labeled (r, c) in Fig. 2).

Segmentation of vegetation typically resulted in some non-vegetative noise pixels incorrectly segmented as vegetation. Morphological dilation is effective in filtering unconnected, noise pixels but is also computationally expensive and thus was not used. Therefore, to reduce the computational burden, a statistical approach based on Otsu's method (Otsu, 1979) was developed to reclassify plant and background regions using the extracted features and minimize the effect of segmentation noise.

After the initial grouping of image rows into background and vegetation regions, Otsu's method (Otsu, 1979) was used to classify vegetation regions as either plants (weeds or maize plants) or noise by determining the optimal threshold to divide the bimodal distribution of each feature. Otsu's method selects a threshold so that the ratio of "between variance" to "within variance" between two classes modes is maximized. For each video segment being analyzed, the histogram of plant region lengths was constructed, and the variance ratio for each possible threshold value was calculated. The threshold that gave the maximum variance ratio was chosen as the optimum threshold. Otsu's method, however, was not independent of amount of noise present and produced a lower threshold when the number of noise pixels was high. Noise pixels result in many small length regions being initially classified as vegetation. The mean vegetation region length was used as the measure of the amount of vegetation regions which were due to noise. Since noise regions were smaller than actual plant regions, the average plant region length would decrease with increasing noise. However, a large average plant region length often indicated the presence of many doubles and triples. The optimal region length depends on plant growth stage and the population. The threshold obtained from Otsu's method was corrected for the amount of noise present using the equation:

$$T_{\text{mod}} = \frac{T_{Otsu}}{\log_{10}(\overline{L}_v)} \tag{2}$$

where T_{mod} is the threshold calculated with noise correction and T_{Otsu} is the threshold calculated from Otsu's method and $\overline{L}_v$ is the mean vegetation length. Noise tended to decrease the mean vegetation region length, so including $\overline{L}_v$ in the denominator of Eqn (2)

counterbalanced the effect of noise on Otsu's method. Otsu's method with noise correction was also used to threshold the background region length and canopy area histograms.

All three features were used to reclassify vegetation and background regions. Vegetation regions with (1) lengths smaller than the plant region threshold and (2) plant areas less than area threshold were reclassified as background regions. Similarly, background regions with lengths smaller than the threshold, were reclassified as vegetation regions. A line was fit to the plant center locations using linear regression. A histogram of the perpendicular distances of the detected plant center locations indicated the assumption of a normal distribution was justified. Second-order statistics were calculated using maximum likelihood estimation (MLE; Hayter, 1996). Any plant with center locations outside the 95% confidence intervals (CI) – called buffer lines – of the estimated crop row line were classified as stray plants or weeds (Fig. 2). The adjusted vegetation and background regions were again compared with threshold values, and misclassified regions were detected and reclassified. The reclassification procedure was repeated iteratively using constant thresholds from the modified Otsu's method operating on the initial classification until the vegetation region count in succeeding iterations changed by less than 5%.

Plant Type Classification

After reclassification of vegetation regions, the region features were recalculated. At this point, each vegetation region contained either one or more weeds or maize plants. The size of weeds and maize plants were assumed normally distributed. However, the number and size of the weeds varied spatially. In the case of no weeds, Otsu's method of thresholding would give an erroneous threshold value because the features would follow unimodal distributions. Therefore, the ratio of "between variance" to "total variance" was used as a measure of separability for two distributions as suggested by Otsu (1979). Separability value ranged from 0 to 1, with a low separability indicating a unimodal distribution indicating a low number of detected weeds compared to maize plants. The threshold obtained from Otsu's method was multiplied by the separability value to obtain a modified threshold that was used to divide the feature histograms into weed and maize regions.

To classify weeds and maize plants, the second-order statistics of the length and area features for both the maize plant and weed distributions were estimated using MLE. If both

the plant length and the canopy area of a vegetation region were less than the 95% CI of the estimated means of those features in the maize plant distribution, then it was classified as a weed. Similarly if a vegetation region had length and area features greater than the 95% CI of their mean in the weed distribution, then that region was reclassified as a maize plant. Once the plants and weeds were classified, any plant which had more than twice the average plant area was considered a double, and thrice the average plant area was considered a triple.

The algorithm was implemented in a Windows application software package using visual C++ and Microsoft Foundation Classes (Microsoft, Redmond, WA). The application used an object-oriented architecture for automated video and GPS capture developed by Shrestha et al. (2003).

Experimental design

The effects of three factors on counting performance were investigated. These factors were: 1) tillage, 2) growth stage, and 3) plant population. Three tillage systems were investigated: "till plant" which did not have tillage prior to planting, "plow" for which a moldboard plow was used resulting in a minimum amount of crop residue on the soil surface, and "spring disk" for which spring tillage was done using a disk or cultivator. Video was collected as the plants varied in growth stages from V3 to V8 (Ritchie et al., 1993). Plant growth stages were classified into three levels namely V3-V4 stages, V5-V6 stages, and V7-V8 stages, to account for the existing variability in growth stages among the plants at the time of data collection. The three levels of population were 39,500, 54,000, and 74,000 plants/ha. The experiment was designed for full factorial interaction resulting in 27 different treatment combinations.

Data collection

Maize row video was collected at the Iowa State University Agronomy and Agricultural Engineering Research Center (Boone, Iowa) during summers of 2001 and 2002. A digital camcorder was mounted on a utility vehicle 0.60 m above the ground with a 0.30 m by 0.40 m field of view. Each captured image size was 480 × 720 pixels with 24 bit color resolution. The vehicle was driven over and parallel to maize rows planted 0.76 m apart with the camera directly over the plants at the speed of about 3.6 km/h. The shutter speed was fixed to 1/1000 s; frames were captured in progressive scan mode; and other camera settings

were set to be automatically adjusted. In the field, the video stream was recorded on a miniDV tape. An RTK-GPS receiver was used to determine the camera location during video recording. The NMEA-0183 standard RMC string transmitted by the GPS receiver was modulated by a VMS 200 unit (Red Hen Systems Inc, Fort Collins, CO), and was recorded on the video tape's audio track (Fig. 1). The maize plants within each experimental unit were also counted manually to compare with automated counted results. Total of 279 experimental units were counted manually during 2001 and 2002.

For spacing analysis validation, the plant stem location from the beginning of each 6.1 m row section to the nearest 1 cm was manually measured and recorded. A steel measuring tape was laid out along the maize row direction, and positions of each plant were recorded. A total of 126 sections were measured manually in the year 2002. Spacing was not recorded for the data from year 2001.

Data analysis

Three way ANOVA was used to study the effects of the main factors (tillage, growth stage, and population) and their interactions on mean plant counting error. For treatments having a significant effect on mean error, treatment means were compared using the Tukey–Kramer multiple comparison method. Homogeneity of error variances across the main factor level was tested using the modified Levene test (Conover et al., 1981). Root mean squared error (RMSE) was calculated and used to represent the standard error of the estimated plant count relative to the manual plant counts. Regression analysis was used to evaluate overall performance of the algorithm with respect to the manual counts.

Plant locations were estimated relative to the end of the crop row section. Each of the plants detected by the algorithm was matched with the nearest plant that was manually measured. The difference in distance along the crop row was calculated and called location estimation error. The SAS General Linear Model procedure was used to test for treatment effects on mean location estimation error and well as error variance with the modified Levene test.

For plant spacing estimates, the interplant distances between every detected plant pair were converted from pixels into physical units. The second order statistics of interplant distances were estimated for both manual measurements and algorithm estimates at each

treatment combination (except for the 74,000 plants/ha population level). Mean manually measured and algorithm estimated interplant distances for each treatment combination were compared using student's t-tests.

Results and discussion

Experimental treatments and video quality

The sensing system could be used over a wide range of daylight conditions. While capturing video of some rows, however, the video camera was mistakenly set to manual aperture, so the camera could not make adjustments when the sky conditions changed resulting in saturated images. This led to saturated pixels in parts of the images resulting in poor plant segmentation performance. Therefore, the 56 experimental units out of 279 experimental units with saturated images were excluded from further analysis. Because of some poor quality video segments, the number of replications over all treatment combinations was not equal and varied from 7 to 11.

A total of 223-paired observations of manual and automated plant counts from 6.1 m (20 ft) row lengths were analyzed using linear regression (Fig. 3). The manual plant counts in an experimental unit varied from 13 to 38 plants, which corresponded to populations of 27, 000 to 81,500 plants/ha. The estimated slope of the regression line was 0.97 which was not significantly different from 1 ($t_{221} = 0.8809$, $P = 0.1191$). The estimated y-intercept was 0.58 which was not significantly different than 0 ($t_{221} = 0.7632$, $P = 0.2821$). The residual plot did not reveal any specific changes in variance across the range of manual counts. The linear model had a coefficient of determination (R^2) of 0.87 and an RMSE of 2.1 plants, which was 8.7% of the mean manual count of 24.1 counts per experimental unit.

This RMSE was somewhat higher than the RMSE reported by Shrestha and Steward (2003) because the algorithm was generalized and tested over a wider range of operating conditions. Much of the error was due to either large weeds or small undetected plants. The addition of other shape features to the weed and plant classification algorithm may be helpful in reducing error.

Plant Count Estimation Error Analysis

Mean Estimation Error

The mean plant count estimation error was significantly different across tillage treatments ($F_{2,216} = 6.71$; $P = 0.0015$). No evidence of significant differences across population ($F_{2,216} = 0.62$, $P = 0.54$) and growth stage ($F_{2,216} = 1.46$, $P = 0.23$) treatments was found. For the spring disk tillage treatment, the algorithm overestimated the number of plants by 0.45 plants per experimental unit. However, for the plow tillage treatment, the automated count was 0.23 plants less than the manual count, and for till plant, the automated count was 1.03 plants less than the manual count (Table 1). Neither spring disk nor till plant tillage systems means were significantly different than plow tillage treatment means, but the mean estimation error was significantly different between spring disk and till plant treatments.

The effect of tillage-growth stage interaction on mean estimation error was significant ($P = 0.0071$; $F = 3.69$), but no evidence of other significant interactions was found. The tillage-growth stage interaction indicates that the tillage system effect depends on growth stage. Further analysis showed that the plant size distribution pattern varied from one growth stage to another. When the plants were small and well spaced, the direction of leaf orientation was random and actual plant length along the row more closely followed a uniform distribution and thus had a larger variance. In addition, when the plants were small, the difference between mean plant size and mean weed size was small. A larger variance and similarity between plant and weed sizes led to misclassification of some larger weeds as plants or some smaller plants as weed. However, at higher growth stages, the plants' leaves were mostly spread out across the row direction; the plant length distribution pattern along the row direction was near normal; and the difference between mean weed and plant size was greater.

The effect of maize plant and weed size on counting error can be explained by examining hypothetical normal distributions of weeds and plants sizes at different growth stages (Fig. 4). Assuming a 95% confidence interval, the region under two normal distributions that overlap is the area of confusion (shaded region in Fig. 4). When maize plant and weed sizes are similar, as in case of the V3-V4 growth stages, the area of confusion is large. When the maize plant size is larger and the maize plant and weed sizes are different,

the area of confusion is smaller as in case of the V5-V6 growth stage (Fig. 4). The average weed size was similar for all maize growth stages, but maize plant size increased substantially across growth stages (Table 2). The expected number of plants or weeds in the area of confusion is proportional to the area of confusion itself. Hence with a larger area of confusion, it is expected that the classification error will be higher. This explanation matched with observations that at earlier growth stages, weeds were the main cause of error in plant estimation and at later growth stages, the overlapped leaves from neighboring plants were the main source of error. Overlapped leaves tended to bias the mean plant length toward higher values. The estimation bias was proportional to the number of plants with leaves overlapping those of neighboring plants (Table 2).

The type and size of weeds varied across tillage treatments. The spring disk tillage treatment had many small weeds adjacent to the maize plants, whereas the till plant system had larger weeds. Larger weeds were similar in size to the maize plants making it difficult to separate weeds from maize plants based on their size. In addition, at the V3-V4 growth stage and low plant population density, plant leaf orientation was nearly random which increased the variance of plant length along the frame row direction. This variance was used to calculate the threshold used to separate plants from weeds. When plant length variance was larger, some of the large weeds were counted as maize plants, and the estimated plant count was biased higher than the manual count.

The amount of residue on the field surface was also different for different tillage systems. For the plow tillage treatment, the field had almost no residue or weeds. There were relatively few noise pixels in the segmented image, and the variance of the plant size distribution was lower. These cleaner surface conditions enabled more accurate estimation of actual plant size and better classification of small weeds from maize plants. For the till plant treatment, however, the field was covered with crop residue and only few weed were visible. Maize plant segmentation was better when the field was covered with residue than when the field had many small weeds. However, with the till plant system, there were many double plants growing close to each other. If there are many double plants in a row, the average plant length estimated by the algorithm was biased towards a higher value. This biasing

caused the double plants to be classified as single plants and led to lower estimates than manual stand counts for till plant.

Error variance

Plant count error variance was significantly different only across growth stages ($F_{2,216} = 15.84$; $P < 0.0001$) and no evidence of differences across tillage treatments and population was found. In particular, the error variance for the V7-V8 growth stage level was significantly higher than the error variance for V3-V4 or V5-V6 growth stages. No evidence of differences in the error variance between V3-V4 and V5-V6 growth stages was found. These results indicate that the uncertainty of population estimates increases as the canopy starts to close in the row. We would expect canopy closure to be a function of both population and growth stage, but for the populations that we had in this data set, no population effect was observed.

A minimum RMSE of 1.04 plants was found for spring disk tillage treatment at V3-V4 growth stage and 40,000 plants/ha. Percentage wise, the RMSE was lowest at 5.2 percent of the mean for the till plant tillage treatment at V3-V4 growth stage and 74,000 plants/ha (Table 3) In general, RMSE was higher for later growth stages. At later growth stages, more plant leaves started overlapping neighboring plants' leaves. This introduced counting error in two different ways. First, the estimated plant length across frame rows was biased to be larger. Second, the larger range of plant sizes increased the variance. These factors increased the area of classification confusion leading to larger error variance. These results indicate overlapped crop plant leaves introduce higher error rates than weeds. However, quantitative relationships between the number and size of weeds and error were not established in this study, as these data were not recorded in the field.

Spacing Accuracy

Plant Location Estimates

No significant effect on the mean plant location estimation error by any of the factors was observed, and the mean error was 1.4 cm. This systematic increase in the estimate of the distance of each plant from start of the row section under analysis was due to the analysis method. Each detected plant was matched with the nearest manually measured plant location. When a manually recorded plant was not detected by the algorithm, the nearest detected plant

was assumed to be the corresponding plant, thus substantially increasing the spacing error (Fig. 5). However, misclassified weeds had no effect on location measurement accuracy, since weeds were not counted during manual measurements. At higher growth stages, more plant canopies were overlapped, and the probability of two plants being counted as one increased leading to an increase in location measurement error.

The variance in the location estimation error was significantly different across growth stages ($F_{3237,2} = 111.03$; $P < 0.0001$), but no evidence of a population effect was found. Error variance increased with increasing growth stages (Table 4) because plant center locations were manually measured differently than they were estimated by the algorithm. Locations were manually measured to plant stems, but the algorithm estimated locations using the median position of each plant region along the crop row. When the plants were smaller in size, the plant leaves were smaller and more symmetrically spread out from the stem. At later growth stages, however, because of the larger, more asymmetrical canopy development, the leaf area centers were more likely to deviate further from the plant stem position, leading to an increased error variance.

Interplant Distance Estimates

A moderately significant tillage effect on location error variance was observed ($F_{3237,2} = 4.46$; $P = 0.0116$) which could be explained by differences in plant spacing uniformity across tillage treatments. A diagram developed by Doerge et al. (2002) was used to visualize the manually measured and estimated interplant spacing. In this diagram, the distance from each plant to its two neighboring plants along the crop row is plotted in a scatter plot (Fig. 6). After plotting the measured interplant distance for three tillage treatments and for a population of 54,000 plants/ha, it was observed that, the number of doubles were higher for the till plant tillage treatment. This result agrees with manual observations from videotapes and was one reason for the underestimation of plant population for this tillage treatment. The estimated plant spacing was also plotted in the same fashion. The manually measured and estimated plant distribution patterns were visually similar.

No evidence of significant differences between the mean measured and estimated interplant spacing was found across combinations of all tillage treatments and two populations (Table 5). However, the modified Levene test for equal variance of manual and

estimated interplant distances showed that the estimated variance was significantly higher than the measured plant spacing standard deviation ($P < .001$ for all cases). The larger variance was primarily due to the algorithm's method of estimating plant center locations.

Conclusions

From this research, it was concluded that the use of plant region statistics is an effective means of classifying early growth stage maize plants for population sensing. Estimation error increases with increases in growth stage, with inter-row canopy closure being a large contributor to error. In addition, these results showed that interplant distances measured in pixels from video frames can be used to effectively estimate interplant distance.

Acknowledgments

This journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 6, was supported by Hatch Act and State of Iowa funds. Additional research support was provided by the Iowa State University Center for Advanced Technology Development. The use of trade names is only meant to provide specific information to the reader and does not constitute endorsement by Iowa State University.

References

Birrell S J; Sudduth K A (1995). Corn population sensor for precision farming. ASAE Paper No. 951334. ASAE, St. Joseph, Michigan, USA

Bullock D G; Bullock D S; Nafziger E D; Doerge T A; Paszkiewicz S R; Carter P R; Peterson T A (1998). Does variable rate seeding of corn pay? Agronomy Journal, **90**(6), 830-836

Bullock D S; Lowenberg-DeBoer J; Swinton S M (2002). Adding value to spatially managed inputs by understanding site-specific yield response. Agricultural Economics, **27**(3), 233-245

Doerge T; Hall T; Gardner D (2002). New research confirms benefits of improved plant spacing in corn. Pioneer Hi-Bred International, Inc. Crop Insights, **12**(2), 1-5

Hayter A J (1996). Probability and statistics for engineers and scientists. PWS Publications, Boston

Nafziger E D. (1996). Effects of missing and two-plant hills on corn grain yield. Journal of Production.Agriculture **9**(2), 238-240

Neilsen R L. (2001). Stand establishment variability in corn. Purdue University, Department of Agronomy Report No. AGRY-91-1

Nielsen R L (1995). Planting speed effects on stand establishment and grain yield of corn. Journal of Production Agriculture **8**(3), 391-393

Otsu N (1979). A threshold selection method from gray - level histograms. IEEE Transactions on Systems, Man, and Cybernetics **9**(1), 62-66

Plattner C E; Hummel J W (1996). Corn plant population sensor for precision agriculture. Proceedings of the Third International Conference on Precision Agriculture. eds. Robert P C; Rust R H; Larson W E, June 23-26, Bloomington, Minnesota, USA. 785-794

Ritchie S W; Hanway J J; Benson G O (1993). How a corn plant develops. Iowa State University Cooperative Extension Service Special report No. 48

Shrestha D S; Steward B L (2003). Automatic corn plant population measurement using machine vision. Transactions of the ASAE **46**(2), 559-565

Shrestha D S; Steward B L; VanWyngarden C; (2003). An object-oriented architecture for field data acquisition, processing and information extraction ASAE Paper No. 033089. ASAE, St. Joseph, Michigan, USA

Shrestha D S; Steward B L; Bartlett E (2001). Segmentation of plant from background using neural network approach. In Intelligent Engineering Systems Through Artificial Neural Networks, Proceedings of Artificial Neural Networks in Engineering (ANNIE) International Conference **11**: 903-908

Sudduth K A; Birrell S J; Krumpelman M J (2000). Field evaluation of a corn population sensor. Proceedings of the Fifth International Conference on Precision Agriculture. eds. Robert P C; Rust R H; Larson W E, July 16-19, Bloomington, Minnesota, USA

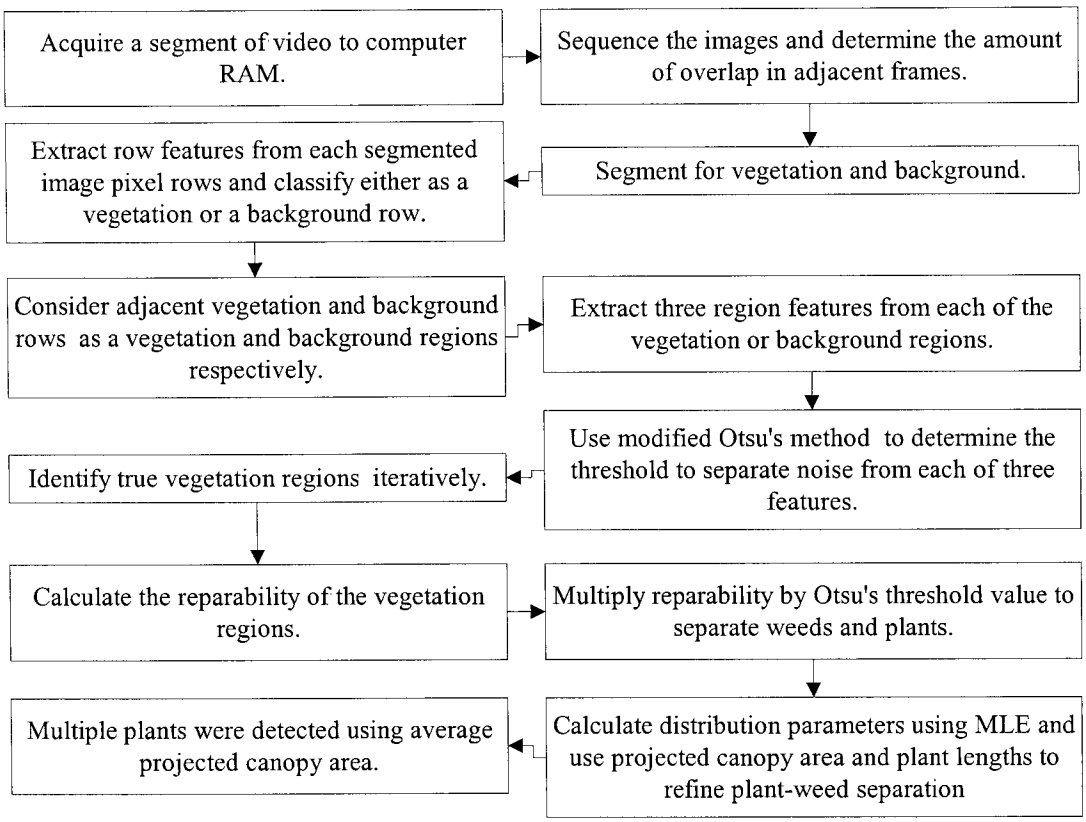


Figure 1. Flow diagram of the algorithm to detect maize plants using machine vision.

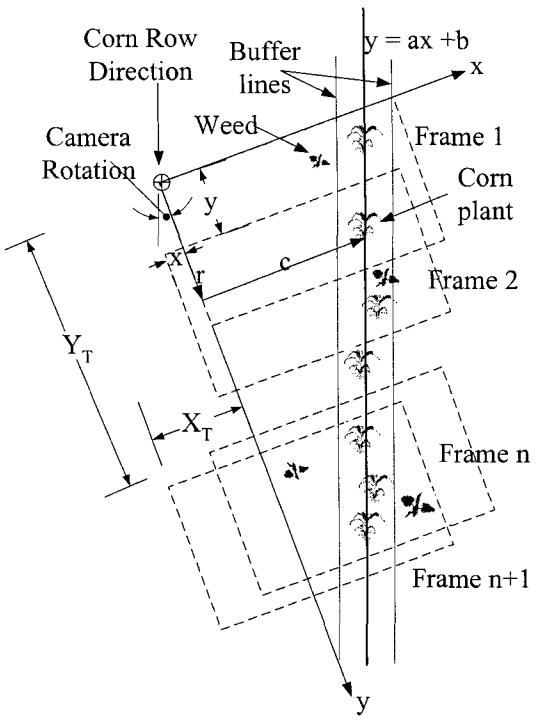


Figure 2. Effect of camera rotation on video frame correspondence. Dashed lines are image frames. The ratio of x/y is constant from frame to frame for camera rotation due to misalignment. This ratio varies from frame to frame for the camera rotation due to vehicle yaw.

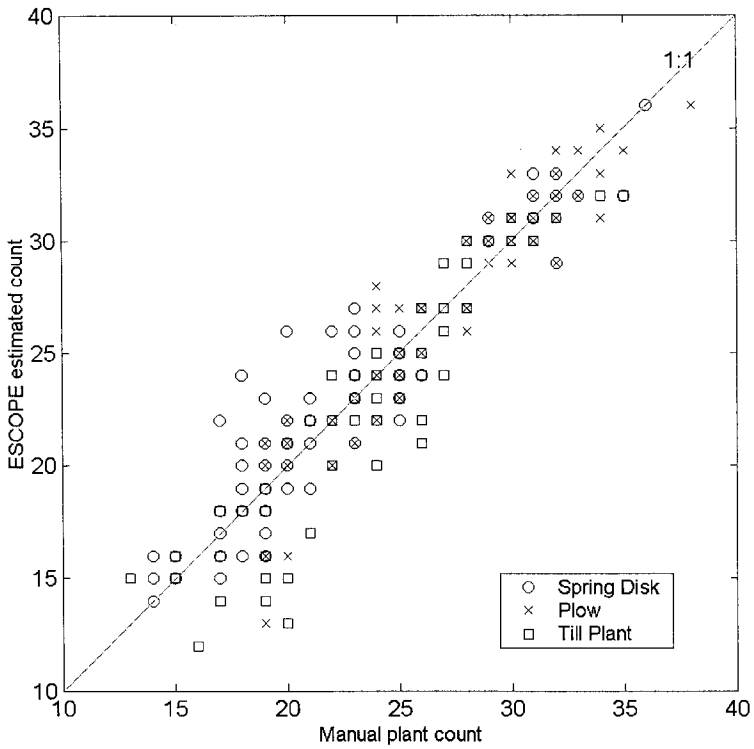


Figure 3. Algorithm estimated maize plant counts related to manual counts for 6.1 m long experimental rows (N = 223) differentiated by tillage treatment.

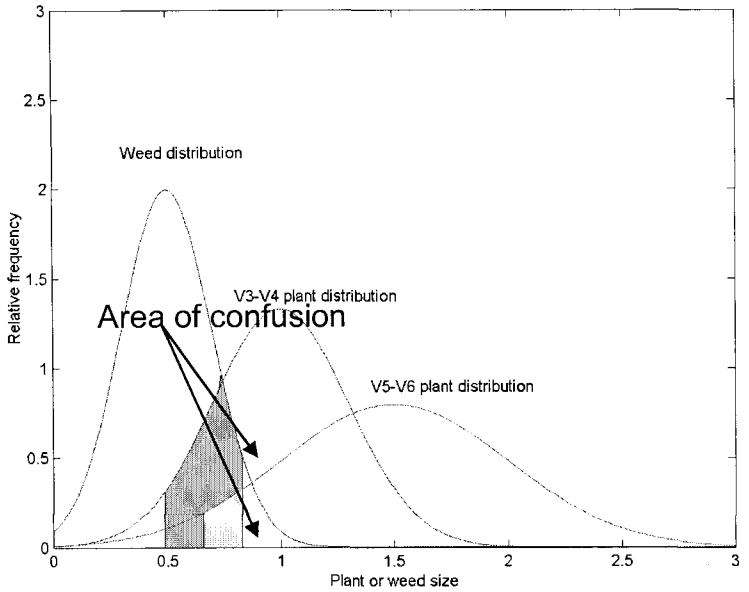


Figure 4. Effects of difference between mean plant size and mean weed size distribution on plant counting accuracy. More the area of confusion, more the expected number of misclassified plants.

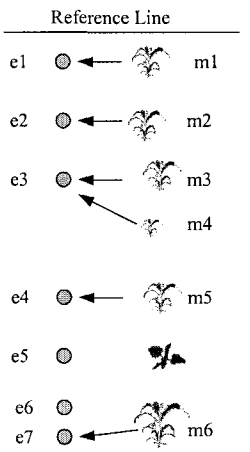


Figure 5. Evaluation of position measurement. m1...m6 are manual measurements of the plants from reference line(Start of the row). e1…e7 are estimated position of plants. m4 was not detected and hence erroneously matched with e3.

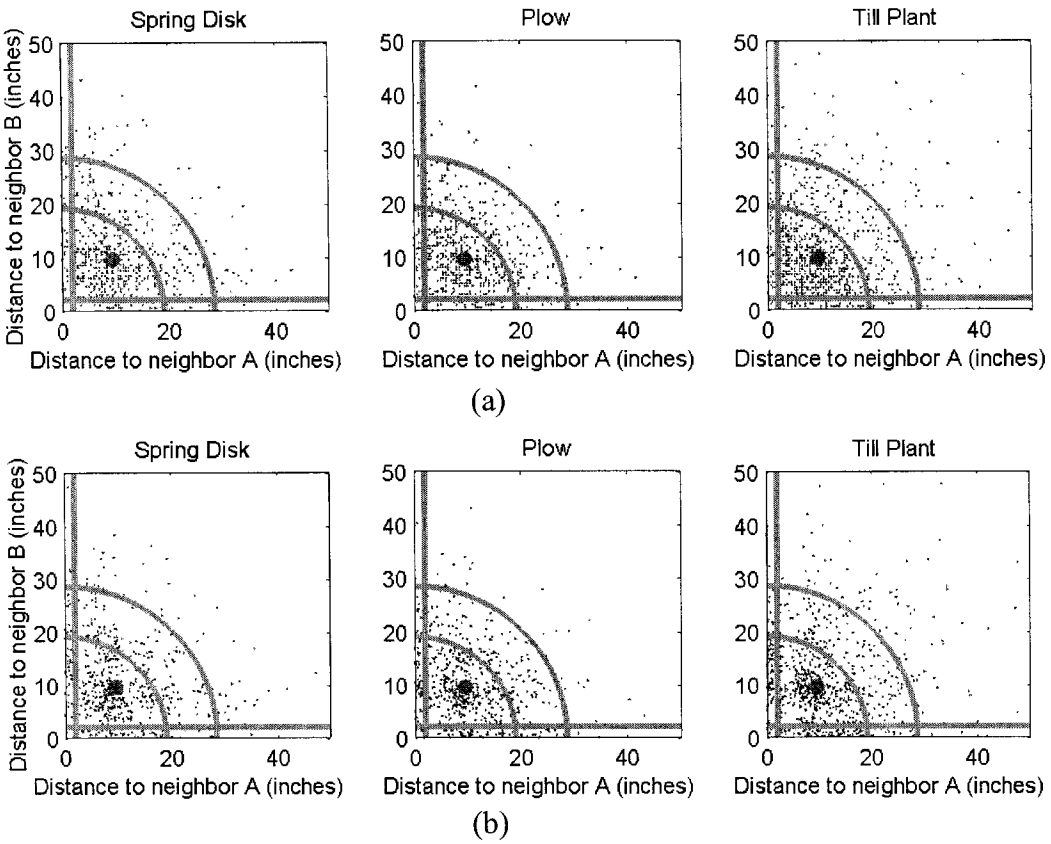


Figure 6. Diagrams indicate the distance from a plant to it two nearest neighbors for (a) manually measured and (b) estimated interplant distance. Points below horizontal line or left of vertical line represent double plants; points nearest origin are triple plants; points between two arcs represent one skipped plant and point outside of the two arcs represent multiple skips

Table 1. Mean error on plant count across different tillage treatment

Tillage	Mean Error
Spring Disk	-0.45a*
Plow	0.26ab
Till Plant	1.03b

*Letters indicate groupings by Tukey-Kramer Test

Table 2. Average plant and weed canopy size and overlapped plants count

Growth stage	Plant size (sq. cm)	Weed size (sq. cm)	Mean number of overlapped plants
V3-V4	18.1 (11.0)*	12.5 (2.6)	1.1
V5-V6	31.8 (16.8)	10.3 (2.3)	2.9
V7-V8	36.25 (23.2)	14.8 (5.8)	7.8

* Numbers in parenthesis is the standard deviation

Table 3. RMSE values for different tillage, growth stage, and plant population density combinations. Number in parenthesis are RMSE as a percentage of mean plant count for a particular treatment combination

Population	Tillage	Growth stage		
		V3-V4	V5-V6	V7-V8
39,500 plants/ha	Spring disk	1.04(5.6)	2.31(12.1)	3.17(17.2)
	Plow	--	--	3.51(18.0)
	Till plant	--	--	3.53(21.1)
54,000 plants/ha	Spring disk	1.39(6.4)	2.30(10.5)	--
	Plow	1.89(8.3)	1.56(6.6)	--
	Till plant	1.96(8.4)	2.19(9.8)	2.62(12.2)
74,000 Plants/ha	Spring disk	1.83(5.8)	--	--
	Plow	2.21(7.0)	1.61(5.2)	--
	Till plant	1.59(5.2)	--	--

-- Data not available

Table 4. Mean distance (cm) between manual plant location and nearest plant counterpart of algorithm estimated plant locations (standard deviation in parenthesis)

Population	Tillage	Growth Stage		
		V3-V4	V5-V6	V7-V8
39,500 Plants/ha	Spring Disk	-0.46 (4.65)	-0.23 (7.62)	1.27 (9.35)
	Plow	--	--	1.55 (10.77)
	Till Plant	--	--	0.76 (12.47)
54,000 Plants/ha	Spring Disk	0.30 (4.55)	-0.05 (6.73)	--
	Plow	-0.10 (4.11)	0.41 (7.29)	--
	Till Plant	-0.05 (4.75)	0.38 (7.24)	0.23 (9.09)

-- Data not available

Table 5. Interplant spacing for different plant population and tillage treatments

Population	Tillage	Interplant spacing (cm.)		t-statistic	P-value
		Measured	Estimated		
16,000 Plants/ Acre	Spring Disk	31.75 (24.64)	31.50 (26.16)	0.095	0.538
	Plow	32.77 (26.42)	32.26 (30.73)	0.131	0.552
	Till Plant	37.08 (28.96)	37.08 (35.05)	-0.038	0.485
22,000 Plants/ Acre	Spring Disk	27.69 (19.81)	27.69 (20.57)	-0.002	0.499
	Plow	26.67 (17.27)	26.67 (19.05)	0.166	0.566
	Till Plant	27.94 (21.59)	27.94 (23.88)	0.054	0.522

[+] Figures in the parenthesis represent standard deviation.

Chapter 6. A Component-Based Software Architecture for Field Data Acquisition, Processing and Information Extraction

A paper submitted to the Transactions of the ASAE.

D. S. Shrestha[1], B. L. Steward[2], C. V. Wyngarden[3]

Abstract

A component-based software architecture was developed to automate site specific field data acquisition, processing, and geo-referenced plant parameter extraction. The architecture supported acquisition and processing of different data streams such as digital video or digital serial communications. Each component in the architecture was developed as a separate module with associated classes. A key component of this architecture was a supervisor class, which communicated and coordinated the operations of all other modules and classes. Standard software components were used in the architecture to do data acquisition. Based on this framework, early stage corn population estimation (ESCOPE) software was developed which grabbed pre-recorded digital video from a vehicle-mounted camera that was passed over corn rows and acquired GPS-NMEA strings which were modulated and recorded on the audio channel. A digital video (DV) capture class was written to acquire video using Microsoft's DirectShow® technology which enabled camera control and video acquisition. Reusability and extensibility characteristics were demonstrated by adding a class to acquire images from the hard drive and also by deriving a new image analyzer class to extract an additional feature. The architecture forms a general framework for developing reusable and extensible software for field data sensing systems.

Keywords: *Precision* agriculture, *machine vision, image processing, COM, DirectX, Unified Modeling Language, field sensing*

---

[1] Graduate Student
[2] Assistant Professor
[3] Undergraduate Student

INTRODUCTION

Precision agriculture (PA) is an important technological development in contemporary agriculture for managing the spatial variability that naturally occurs in crop production (Schueller et al., 2002). The National Research Council (1997) refers to PA as a management strategy that uses information technologies to bring data from multiple sources to bear on decisions associated with crop production. The key idea behind PA is to measure and manage spatial variability to optimize the crop production system. Spatial variability can be categorized into yield, field, soil, crop, and management variability.

Yield variability is a function of other four variabilities. Field, soil, and crop variability can be assessed with direct or indirect measurements, and models have been developed to explain yield variability in terms other variabilities. However, most crop models account for a small number of factors that may limit yields in the field (Gary et al., 1998; Irmak et al., 2001; Schueller et al., 2002). Currently no complete precision agriculture system exists; rather various components of traditional crop management systems have been evaluated separately regarding their potential for site-specific management (Pierce and Nowak, 1999). To face the challenge of diversity, modelers will need to adopt more generic approaches (Gary et al., 1998). A major reason why researchers have only included a small set of variables in their model was that the characterization of field-level variability had been generally inaccessible or prohibitively expensive to acquire. Advances in electronics, communications, and software over the past several decades have removed those earlier impediments. Inexpensive sensors and microprocessors coupled with integrating software now enable agricultural producers to collect vast amounts of geo-referenced data (Schmoldt, 2001).

Different sensors have been developed to measure field scale variability in several crop production parameters. Yield sensors for major crops are approaching maturity and are commercially available (Zhang et al., 2002). Sensors have been developed for measuring soil properties such as soil organic matter and moisture content (Hummel et al., 2001), electrical conductivity (Lund et al., 2000) and nutrients (Birrell and Hummel, 2000). Sensors have also been developed to measure plant parameters such as corn population (Birrell and Sudduth, 1995), nitrogen status (Goel et al., 2003) and leaf area index (Johnson et al., 2003). While not

yet deployed on a field scale, image-based crop growth measurement has been shown to be effective in measuring and modeling crop plant growth in laboratory or greenhouse applications (Morden et al., 1997; Tarbell et al., 1991; Tarbell and Reid, 1991; Van Henten and Bontsema, 1995). Machine vision-based algorithms have been developed to estimate several crop and field parameters such as weed infestations (El-Faki et al., 2000; Tang et al., 2000), plant shape and size (Nishiwaki et al., 2001), plant population (Shrestha and Steward, 2003) and plant height (Shrestha et al., 2002).

Bottlenecks in successful application of PA include a lack of (1) developed sensing technologies needed to adequately characterize field-scale spatial variability, (2) flexible data acquisition and processing systems that can be deployed in a field to gather and process data, and (3) agronomic knowledge relating crop inputs and plant response to those inputs. This work focuses on the second of these bottlenecks through the development of a flexible software architecture for systems that can acquire geo-referenced field data and extract plant growth and soil parameters from these data. Such an architecture should enable continued system development and reusability when new algorithms are added to the software.

Developing a software system architecture prior to its construction or renovation is as essential as having a blueprint for a large building. There are many advantages to developing system architecture. First, it breaks down the entire task into individual standalone modules so that an individual or small group of individuals can work on a component. Second, it offers ease in management of individual components. Third, architecture facilitates reuse, modification and improvement of the software. Fourth, the architecture helps isolate errors and allows individual module testing. Fifth, it improves readability and brings clarity.

A growing body of literature illustrates the utility of object-oriented architectures in agricultural applications. Object-oriented architectures are being investigated in the area of crop growth modeling (Beck et al., 2003; Pan et al., 2000). The development of agricultural autonomous vehicles is being facilitated through the use of object-oriented architectures that help enhance the collaboration of non-collocated research groups (Blackmore et al., 2002; Sorensen et al., 2002; Torrie et al., 2002; Will et al., 2002). These architectures enable reuse of software components and enable design abstraction and systematic thinking about highly complex systems.

The objective of this research was to develop a software architecture for a field data acquisition and processing system which can receive data from several sources, process these data, and output desired parameter estimates. Based on this architecture, an early stage corn population estimation (ESCOPE) system was developed. Using ESCOPE as a case study, following characteristics of the architecture were demonstrated:

1. Expansibility of data acquisition from different input data channels through the use of standard hardware interfaces.
2. Reusability of the developed algorithm.
3. Extensibility through ease in incorporating new data processing algorithms.

This architecture is expected to be particularly beneficial for several different groups of people. The first group consists of researchers who would like to measure crop growth parameters such as crop greenness, plant height, or top projected canopy area. The benefit for this group of people comes from reducing the distractions typically encountered when introducing new information streams or processing algorithms into a system. The architecture aids system development as it is designed to use a standard hardware interface. Another group is crop modelers who need different field level variability data to calibrate models, develop input prescriptions, and assess economic and environmental risks. As PA is further developed, crop consultants and agricultural producers will benefit from such a system as a component within a larger decision support system.

System Architecture

An component based object-oriented programming (OOP) approach was followed because of its many advantages over structural programming. In general, OOP uses software objects that resemble real world objects which are easier to understand and conceptualize. OOP also enables the reuse of an object and easy modification through class inheritance and is less prone to error because of its capability to encapsulate data. Object-oriented systems can be effectively described using Unified Modeling Language (UML). UML is a modeling tool for specifying, visualizing, constructing, and documenting the artifacts of software systems (Object Management Group, 2003). UML is widely accepted by the software engineering community for constructing and analyzing the early stages of system

development (Aleman and Alvarez, 2000). The architecture can be described using three different UML diagrams –context, use case, and class diagrams.

Context Diagram

The chronological activity of the entire system is shown in a context diagram which shows the system boundary and system's role in a bigger context (Staugaard, 2001). A field data acquisition, processing and information extraction system must first acquire field data through a computer input port receiving a signal (fig. 1). Examples of potential signals include an analog or digital signal from a soil temperature sensor or a video signal from a video camera, or a National Marine Electronics Association (NMEA) string produced by the GPS receivers. Whatever the signal, the architecture is general enough to develop systems that have the ability to accept, analyze, and correlate data from those signals.

Processing of stored data can be done in two different modes: a batch mode and continuous mode. In the batch mode, data are acquired and stored in memory and then processed. The signal source is paused during data processing. When processing is completed, computer memory is freed and another segment of data is acquired. With this mode, the sequence of events is well defined and processing does not need to be confined to real time constraints. In the continuous mode, the data acquisition and processing occur in parallel. The data is read into a circular buffer. Once a predetermined portion of the buffer is filled, data processing starts while data acquisition continues simultaneously. This mode requires that processing is done within real time constraints.

Resulting from data acquisition and processing are geo-referenced site or plant-specific parameter estimates which are passed onto the storage system (fig. 1). The extracted information may be used in a crop model to establish relationships between yields and a set of indicator variables, for decision support in making crop production management decisions, or to produce a GIS database which archives field spatial variability information.

Use-case Diagram

A use-case diagram attempts to graphically depict the system users – called actors – and actions – called use-cases. This diagram shows the role each actor will play in interacting with the system. The main actors of the system are digital data, video and GPS signals, researchers and management. In a typical field application scenario, the system acquires

different data (fig. 2); researchers and farm managers then decide on what kind of field parameters to extract. Choices about the parameters to extract depend on the type of field data and processing capabilities. Researchers may also incorporate new processing modules into the system. This architecture serves as a container that incorporates new modules and combines the new output along with old outputs. Crop production management can make field-management decisions based on the extracted information.

Class Diagram

Class diagrams are used to show the relationship among systems objects. A group of activities pertaining to some specific task was identified from the use-case diagram as a module consisting of one or more classes. Three distinct modules were identified: 1) the data capture module 2) the data analysis module, and 3) the display module.

In addition to the modules, a Supervisor class was designed to coordinate the interaction of these modules. The Supervisor class keeps track of all activities in the system and synchronizes different processes. This class creates all other classes as needed to handle windows and internal messaging. The Supervisor class also holds pointers to data stored in the memory (fig. 3). The data capture module is comprised of classes that deliver data from multiple input streams to the supervisor class. Each input stream requires a corresponding data capture module that is independent of other data capture modules. This independence enables the supervisor class to select any combination of input streams. Every data capture module is required to control its data source (if processing in the batch mode), acquire data, perform error checking and format the data so that it can be passed to the supervisor. For a machine vision-based architecture with geo-referencing, at least two data capture modules are required: a video capture module and a GPS capture module. Each aggregation relation to supervisor class represents a separate module (Fig. 3).

Video Capture Module

The video capture module captures a video stream and checks for any errors in the data captured. This module consists of the DVcapture, Grabber, and DVIterator classes.

DVCapture Class

Component-based software development has enabled the reusability of program code. Reusable code can be used, without modification, to perform a specific services regardless of what application uses the code. Programmers can develop one piece of software that uses services provided by some other software without having to understand the details of that software. In addition, it is much easier and safer to use standard components than writing a new program to do the same task. Because of these advantages, the DVCapture class was developed taking advantage of Microsoft DirectShow® technology which is based on the component object model (COM).

Microsoft® COM technology is a widely-used software model which provides a standard protocol for object intercommunication and reusability (Root and Boer, 1999). COM is a platform-independent object-oriented system for creating software components that have defined interfaces for interaction. COM objects can be created with a variety of programming languages and can be embedded in programs that are written in different languages (MSDN Documentation, 2003b). DirectX® (Microsoft Corp., Redmond, WA) consists of a group of COM objects, which provides a standard development platform for Windows-based PCs by enabling software developers to access specialized hardware features without having to write hardware-specific code. DirectShow is one of the member components of DirectX. DirectShow technology can be used to control video cameras for playback and video image acquisition from various sources and formats.

The DVCapture class creates the Grabber COM object so that camera control functions can be called. A filter graph, which is specialized class for video parsing, is created in this class to capture video frames. Once the filter was built, different functions were written so that more intuitive camera commands like "Start Capture" and "Stop Capture" could be used to control the camera from the Supervisor class. Some of the important tasks this class performs are 1) finding video input devices, 2) watching for events such as end of tape or external stopping of the camera, 3) getting video streams and splitting them into frames, 4) playing, pausing, and stopping the camera, and seeking to a location.

Grabber Class

The Grabber class is derived from the DirectX base class, CUnknown, and the ISampleGrabberCB interface. This class was designed for communicating with a digital video player. Since this class is derived from CUnknown, it is a COM object by itself. This strategy allows some call back methods of ISampleGrabberCB to be overwritten for user-defined memory management. It also enables acquisition of the absolute video track number at a regular interval and comparison of the track number with actual number of frames acquired so that frame drops can be detected.

DVIterator Class

The DVIterator class is a container class for captured image frames. On initialization, the Grabber class creates this class and passes a pointer from this class to the DVCapture class. In turn, the DVCapture class passes back the pointer to the Supervisor class, so that the Supervisor class can pass it to any class that requests the image storage location. This class keeps track of the image frames in computer RAM, so that the frames can be retrieved in sequence for processing. In addition, once the processing is completed, this class frees memory for storing new data. This class also 1) checks for valid data 2) moves a pointer to the first or last image frame, 3) accesses the frame from a specified location, and 4) replaces a frame at a specified location.

GPS capture module

The GPS capture module is similar to the video capture module except that this module is designed to capture NMEA strings from the serial port. The module consists of PortUtility, GPSCapture, and GPSIterator classes.

PortUtility class

The PortUtility class initializes serial port settings. In this architecture, the software is expected to read the digital data like GPS and other measurements through serial ports such RS-232 and CAN ports. This class reads data for a specified time period or within specified time boundaries. This class relieves programmers from writing their own serial port setting class repeatedly for each channel.

GPSCapture class

The GPSCapture class is derived from the PortUtility class to capture the GPS data from the serial port. This class creates a processing thread to start and stop capture in the time specified by Supervisor class. This class also parses NMEA strings to obtain GPS parameters or measurements requested by the Supervisor class. The GPSCapture class contains the GPSIterator class to store the GPS data and derived information.

GPSIterator Class

The GPSIterator class is a container class for storing GPS data. This class is aggregated in the GPSCapture Class, which holds a pointer to this class and passes this pointer back to the Supervisor class so that the Supervisor class can pass it to the classes that act on GPS data. The GPSIterator class provides functionalities to step through the acquired GPS data in RAM. This class also validates acquired data and informs the Supervisor class about any errors it encounters. The function of this class is similar to the DVIterator class except that the class handles GPS data rather than video data.

Data Analysis Module

The application's core image processing algorithms reside in the data analysis module. The supervisor delivers the data pointer to this module along with other information such as previous analysis results to improve the accuracy and speed of the analysis algorithms. The data analysis module returns the extracted plant and soil parameters to the supervisor. The number of parameters may vary and are processing algorithm dependent. Data analysis modules are specific to the application and many have one or more classes. However, the communication format between an analysis class and supervisor class is fixed by the architecture.

Display Module

The display module presents results to the application users. The supervisor passes to this module a pointer to the results table. The results are stored in the memory in a tabular form with each row containing the information for a particular plant. This module contains the classes for graphical display and storage of analysis output. The type of the graphics may vary from one application to another.

Supervisor Class

The Supervisor class is the main controller of all classes. It initiates, controls, and releases all other classes as needed. This class holds a pointer to each of the classes it creates and keeps track of all the events of the application. The Supervisor class was derived from the Microsoft Foundation Class (MFC) CDocument class (MSDN Documentation, 2003a) and thus inherits many functionalities of the CDocument class to manage the application data and share the data with the display module.

Key to the architecture's effectiveness are the interfaces between the modules and the supervisor class. Well-defined interfaces enable "plug-and-play" component-based software architecture, which allows easy upgrades to the application. When a new algorithm is developed to analyze data, for example, a new data analysis module replaces the old module. As long as the new module conforms to the defined interface, the application can easily be modified to take advantage of the new algorithm. The interface consists of three types of messages 1) error messages, 2) update messages and 3) state messages. Error messages containing the cause of the error are passed when anomalous situations are encountered. Update messages are passed from the data analysis module to update the supervisor class with new parameter values. State message inform the supervisor class of the state of the different modules. The message handling functions are made virtual so that if a new class is added in the architecture which needs a different message handling method, a new Supervisor class can be derived from the old Supervisor class to overwrite the message handling functions.

On startup, the Supervisor class creates relevant capture classes that check hardware readiness. If hardware is not ready, the capture class associated with that hardware sends an error message, and the supervisor class redirects the message to the display module which displays a warning message. Similarly, when the user turns on or off any of the data input channels, the supervisor class communicates with other classes through the messaging system.

The Supervisor class is strictly message driven. The actions taken by the supervisor class depends on type of message it received from other modules and the operating system. While acquiring video from a camcorder, for example, the DVCapture class may send the

end of tape message to the Supervisor class. Then the Supervisor class calls the function pertaining to that specific message.

When a new data processing algorithm is added to the system, the following steps need to be performed:

1. Decide on data channels needed and list those channels in the Supervisor class.
2. Set the message type and handling function in the Supervisor class for the additional extracted parameters.
3. Add an entry in the output database structure for new parameters.

With these three steps, future researchers can add any number of data processing algorithms.

Case Study: ESCOPE

Based on the architecture described above, software was developed to measure corn plant population and spacing with machine vision. Software development was done using Visual C++ (Microsoft, Redmond, WA). ESCOPE acquired video and GPS data, processed this data, and displayed results to the user. As a case study, ESCOPE illustrates the reusability and extensibility facilitated by the architecture.

Video acquisition and processing

Video was recorded in the field and then brought to the laboratory for analysis. Digital video signal was transferred to the computer through an IEEE 1394 (Fire Wire) connection, whereas GPS NMEA strings were transferred using the serial port. Video frames were spatially correlated with GPS data.

Image processing was accomplished by the ImageAnalyzer class which operated on the video frames stored in RAM. The Supervisor class provided a pointer to the DVIterator class and the GPSIterator class to the ImageAnalyzer class. The ImageAnalyzer class was designed to run on its own thread so that during real time analysis, the processing and capturing could be done in parallel.

The ImageAnalyzer class extracted the memory location of each video frame from the DVIterator class and mosaicked each frame discarding the portion of each frame that was overlapped with the previous frames. After mosaicing the images, two features were extracted from each row of the composite image: the total number of plant pixels, and their

median position. Adjacent rows of the same class were grouped together and iteratively refined for final plant counting. The details of the image processing were described in Shrestha and Steward (2003). The ImageAnalyzer class not only detected the plants, but also detected the field markers, and reported the detected plant center and marker center locations in terms of pixels counted from the beginning of the video segment to the Supervisor class.

The video time code was used to synchronize video images with the GPS data. Video time code was associated with each image frame and read independently through both the FireWire and RS-232 ports. This allowed the association of GPS signals with each video frame and provided a basis for geo-referencing each plant detected through image processing. However, the GPS signal update rate was lower than the frame rate; hence the GPS locations for intermediate frames were linearly interpolated. Once the Supervisor class received updates from the ImageAnalyzer class, it interpolated the location of each plant for its GPS location and wrote to the output database table.

Information display and output

A plant-centric database format was used to output the information. In the plant-centric database, other data such as GPS location are interpolated for each detected plant. A plant-centric database is more suitable where information about individual plants is more important than statistics over a region like in the study of plant spacing and canopy area measurement.

A display module was written to display the extracted information. Once a video segment was processed, the Supervisor class sent a pointer to the new parameters and a message to the display module to update the user interface. The display module produced three different windows panes to display different information. The first pane displayed the composite images (fig. 4) which permits visual evaluation of image correspondence and plant detection performance. The second window pane displayed the plant population in the form of a bar graph along with numerical displays of maximum, minimum, and current plant densities (fig. 4). The third pane displayed processing and other information in text format which was used during development. Finally, the program wrote the parameter for every detected plant to a file in a tabular format (fig. 5) which includes the plant number, the plant location in terms of number of pixels from the beginning of a video segment, the plot

number, the video time code when that particular plant was found, and the interpolated latitude and longitude of the corn plant.

Discussion

The object-oriented architecture provided a framework for developing a field video acquisition, image processing, and plant-centric parameter extraction software. Camera control and video acquisition were successfully implemented using Direct Show®. Use of the architecture led to more intelligent software development and efficient data flow and processing. By using industry-standard hardware interface components, the architecture facilitated data channel expansion.

In order to demonstrate component reusability, after coding the program for acquiring the video signal from the camera, a additional class called FileCapture was developed and incorporated into the program to read a sequence of video frame images from the hard drive. The FileCapture class was derived from the Grabber class and only the supervisor class needed to be modified to incorporate the FileCapture class. Hence the Grabber class and other classes were reused to add this file capture functionality. The FileCapture class also used DirectShow® and has similar functionality as DVCapture except for camera control.

To test the extensibility of the program, a class, ImageAnalyzer2, was derived from the ImageAnalyzer class with the functionality needed to estimate the top projected canopy area (TPCA) of each plant. The newly derived class had all the functionalities of the original ImageAnalyzer and provided updates of TPCA to the supervisor class. To accommodate this new feature, only the database structure in the Supervisor class and output format needed to be changed. When the video frames were read from the hard drive using the FileCapture class and processed with the new ImageAnalyzer class, the output table contained different information than the original table (fig. 6). The architecture facilitates ease in extending image analysis to extract new parameters.

Conclusion

Component-based object-oriented architecture developed in this work addresses a need in precision agriculture by providing a general framework for the development of field data acquisition, processing and information extraction systems. Industry-standard hardware

interface components provided the basis for independent data capture modules that could be replicated for data channel expansibility. The addition of new features demonstrated the reusability and extensibility of the architecture since only the supervisor class needed to be modified when incorporating a new image processing class.

Acknowledgements

This journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 4003, was supported by Hatch Act and State of Iowa funds. Additional research support was provided by the Iowa State University Center for Advanced Technology Development.

References

Aleman, J. L. F. and A. T. Alvarez. 2000. Can intuition become rigorous? Foundations for UML model verification tools. In *11th International Symposium on Software Reliability Engineering* 344-355. IEEE.

Beck, H. W., P. Papajorgji, R. Braga, C. Porter, and J. W. Jones. 2003. Object-oriented approach to crop modeling: concepts and issues. In *Proc. World Congress on Computers In Agriculture and Natural Resources* 297-305. F. Zazueta and J. Xin, eds. St. Joseph, Mich.: ASAE.

Birrell, S. J. and J. W. Hummel. 2000. Membrane selection and ISFET configuration evaluation for soil nitrate sensing. *Transactions of the ASAE* 43(2): 197-206.

Birrell, S. J. and K. A. Sudduth. 1995. Corn population sensor for precision farming *ASAE Paper No. 951334*. St. Joseph, Mich.: ASAE.

Blackmore, S., S. Fountas, and H. Have. 2002. A proposed system architecture to enable behavioral control of an autonomous tractor. In *Proceedings of automation technology for off-road equipment* 13-23. Q.Zhang, eds. Chicago, IL. St. Joseph Mich.: ASAE.

El-Faki, M. S., N. Zhang, and D. E. Peterson. 2000. Weed detection using color machine vision. *Transactions of the ASAE* 43(6): 1969-1978.

Gary, C., J. W. Jones, and M. Tchamitchian. 1998. Crop modeling in horticulture: state of the art. *Scientia Horticulturae* 74(1-2): 3-20.

Goel, P. K., S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau. 2003. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Computers and Electronics in Agriculture* 39(2): 67-93.

Hummel, J. W., K. A. Sudduth, and S. E. Hollinger. 2001. Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. *Computers and Electronics in Agriculture* 32(2): 149-165.

Irmak, A., J. W. Jones, W. D. Batchlor, and J. O. Paz. 2001. Estimating spatially variable soil properties for application of crop models in precision farming. *Transactions of the ASAE* 44(5): 1343-1353.

Johnson, L. F., D. E. Roczen, S. K. Youkhana, R. R. Nemani, and D. F. Bosch. 2003. Mapping vineyard leaf area with multispectral satellite imagery. *Computers and Electronics in Agriculture* 38(1): 33-44.

Lund, E. D., C. D. Christy, and P. E. Drummond. 2000. Using yield and soil electrical conductivity (EC) maps to derive crop production performance information. In *Proceedings of Fifth International Conference on Precision Agriculture,* P.C.Robert, R.H.Rust, and W.E.Larson, eds. Bloomington, MN, USA. Precision Agriculture Center, University Of Minnesota.

Morden, R. E., P. P. Ling, and G. A. Giacomelli. 1997. An automated plant growth monitoring system using machine vision. *ASAE Paper No. 974033.* St. Joseph, Mich.: ASAE.

MSDN Documentation. 2003a. CDocument. Microsoft Inc. Available at: http://msdn.microsoft.com/library/default.asp?url=/library/en-us/wcemfc/htm/cdocumnt_2.asp. Accessed on 29 May 2003.

MSDN Documentation. 2003b. The Component Object Model. Microsoft Inc. Available at: http://msdn.microsoft.com/library/default.asp?url=/library/en-us/com/htm/com_757w.asp. Accessed on 25 Apr. 2003.

National Research Council. 1997. Geospatial and information technologies in crop management. National Academy Press. Washington.

Nishiwaki, K., T. Togashi, K. Amaha, and K. Matsuo. 2001. Estimate crop position using template matching in rice production. *ASAE Paper no. 013103*. St. Joseph, Mich.: ASAE.

Object Management Group. 2003. Unified modeling language specification, Ver.1.5. OMG object management group. Available at: http://www.omg.org/technology/documents/formal/uml.htm. Accessed on 25 Apr. 2003.

Pan, X., J. D. Hesketh, and M. D. Huck. 2000. OWSimu: an object-oriented and Web-based simulator for plant growth. *Agricultural Systems* 63(1): 33-47.

Pierce, F. J. and P. Nowak. 1999. Aspects of precision agriculture. *Advances in Agronomy* 67: 1-85.

Root, M. D. and J. R. Boer. 1999. *DirectX Complete.* Ch. 1, 3-18. NY:McGraw-Hill.

Schmoldt, D. L. 2001. Precision agriculture and information technology. *Computers and Electronics in Agriculture* 30(1-3): 5-7.

Schueller, J. K., J. D. Whitney, T. A. Wheaton, and L. A. Balastreire. 2002. Accuracy and data characterization analysis in precision agriculture using matlab. *ASAE Paper No. 021044.* St. Joseph, Mich.: ASAE.

Shrestha, D. S. and B. L. Steward. 2003. Automatic corn plant population measurement using machine vision. *Transactions of the ASAE* 46(2): 559-565.

Shrestha, D.S., Steward, B.L., Birrell, S.J., and Kaspar, T.C. 2002. Corn Plant Height Estimation Using Two Sensing Systems. *ASAE Paper No. 021197.* St. Joseph, Mich.: ASAE.

Sorensen, C. G., H. J. Olsen, A. P. Ravn, and P. Makowski. 2002. Planning and operation of an autonomous vehicle for weed inspection *ASAE Paper No. 021177.* St. Joseph, Mich.: ASAE.

Staugaard, A. C. 2001.The object-oriented paradigm using UML. In Structured and object oriented problem solving using C++.Upper Saddle River, NJ:Prentice Hall.

Tang, L., L. Tian, and B. L. Steward. 2000. Color image segmentation with genetic algorithm for in field weed sensing. *Transactions of the ASAE* 43(4): 1019-1027.

Tarbell, K. A. and J. F. Reid. 1991. A computer vision system for characterizing corn growth and development. *Transactions of the ASAE* 34(5): 2245-2255.

Tarbell, K. A., D. K. Tcheng, and J. F. Reid. 1991. Corn growth and development attributes obtained using inductive learning techniques. *Transactions of the ASAE* 34(5): 2264-2271.

Torrie, M. W., D. L. Cripps, and J. P. Swensen. 2002. Joint architecture for unmanned ground systems(JAUGS) applied to autonomous agricultural vehicles. In *Proceedings of Automation Technology for Off-road Equipment Conference* 1-12. Q.Zhang, eds. Chicago, IL.

Van Henten, E. J. and J. Bontsema. 1995. Non-destructive crop measurements by image processing for crop growth control. *J.Agric.Engng Res.* 61: 97-105.

Will, J. D., J. F. Reid, N. Noguchi, and Q. Zhang. 2002. Software architecture design for automation of off-road equipment. In *Proceedings of Automation Technology for Off-road Equipment* Q.Zhang, eds. Chicago, IL.

Zhang, N. Q., M. H. Wang, and N. Wang. 2002. Precision agriculture - a worldwide overview. *Computers and Electronics in Agriculture* 36(2-3): 113-132.

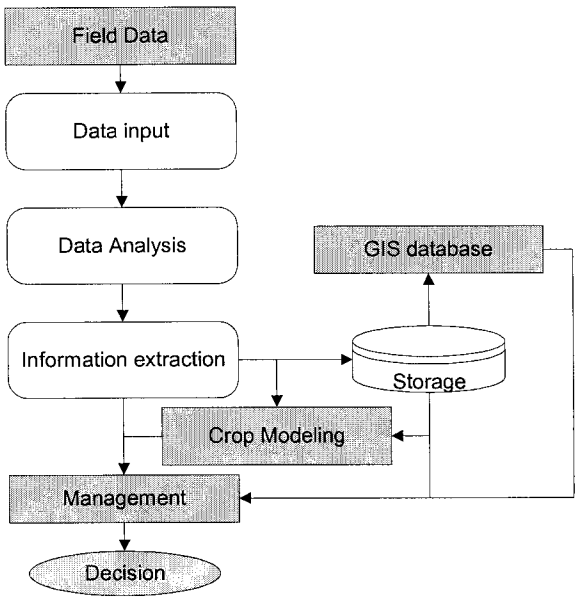


Figure 1. Context diagram for field data acquisition, processing and information extraction system. The software architecture developed acquires recorded data, analyzes, produces site-specific information, and stores into hard drive (Non-grayed boxes).

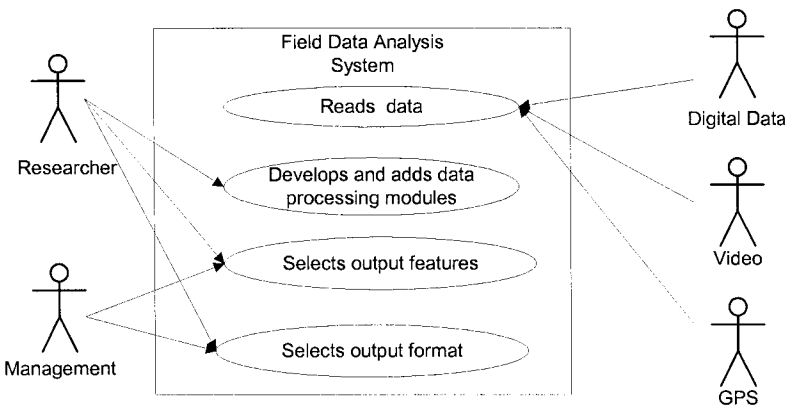


Figure 2. Use-case diagram. The field data analysis system interacts primarily with these actors: digital data, Video, GPS, researcher and management.

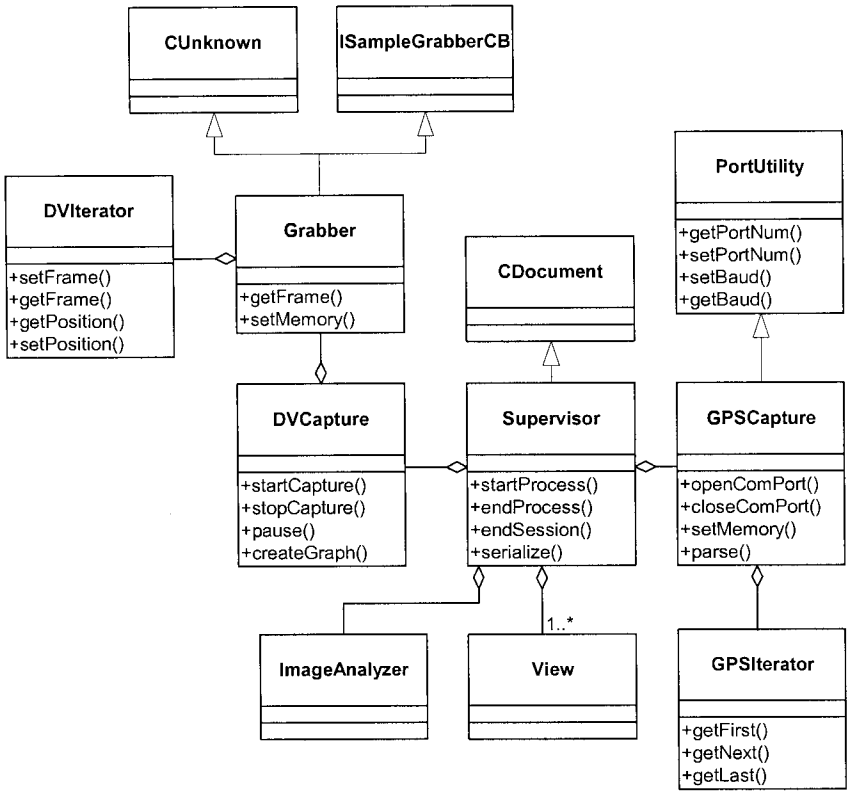


Figure 3. Class diagram. Supervisor class is the heart of the structure, which manages all other classes and activities.

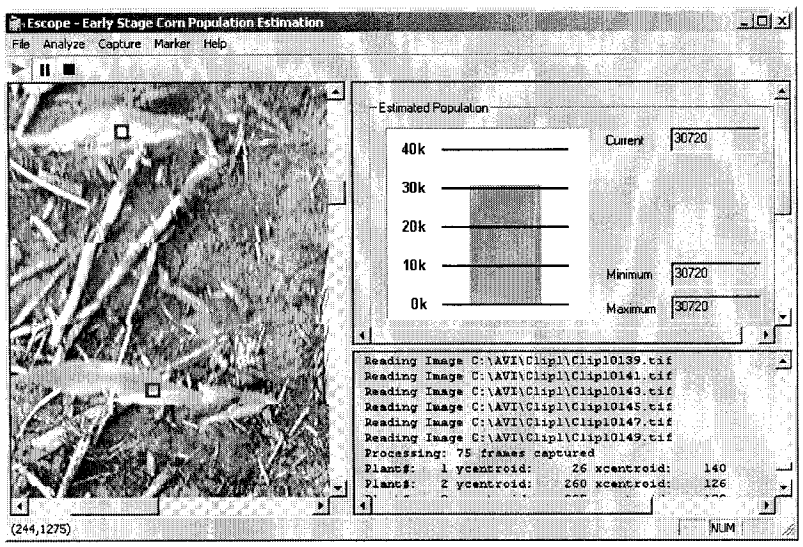


Figure 4. Interface of the program developed to estimate the plant location and spacing measurement. The program was named ESCOPE for early stage corn population estimation.

Plant#	x	y	Plot#	Timecode	Latitude	Longitude
1	126	602	1	00:09:26:14	42.01752513	-93.76642624
2	233	821	1	00:09:26:20	42.01752513	-93.76642823
3	254	1082	1	00:09:26:26	42.01752513	-93.76643022
4	306	1400	1	00:09:27:06	42.01752513	-93.76643354
5	259	1635	1	00:09:27:12	42.01752512	-93.76643553
6	581	2011	1	00:09:27:22	42.01752512	-93.76643885

Figure 5. Plant output of ESCOPE. First column shows plant detected. x and y shows plant position from the top of the row in pixels. Time code is the video time when plants were found. Latitude and Longitude are estimated GPS location of the plant.

Plant#	x(in)	y(in)	Plot#	Area(sq.in)	Image#
1	4.0	0.7	1	4.3	1
2	4.8	8.4	1	3.9	5
3	4.7	15.5	1	3.0	9
4	5.9	23.3	1	2.7	17
5	4.8	32.6	1	2.9	27

Figure 6. Output from modified ESCOPE. A FileCapture class read the images sequence from hard drive without GPS and new ImageAnalyzer calculated top projected canopy area of each plant.

Chapter 7. Corn Plant Height Estimation Using Two Sensing Systems

Presented at ASAE annual international meeting, 2002. ASAE paper No: 021197

Dev Sagar Shrestha[1], Brian L. Steward[2], Stuart J. Birrell[2], Thomas C. Kaspar[3]

Abstract

Crop growth rate is influenced by climate, soil properties, and agricultural inputs. Information on the spatial patterns of crop growth in the field can potentially be used to manage agricultural inputs. Two different sensing approaches, stereo vision and ultrasonic, were investigated as candidate technologies for vehicle-based corn height sensors. For the stereo vision method, a chain code-based stereo correspondence technique was developed to determine the disparity in the stereo image pair. Images were taken using one camera from a series of precisely controlled locations to generate the stereo effect. The ultrasonic sensor measured the distance to an object by detecting the time delay of ultrasonic echoes. The echoes from the plant canopy were recorded, and collar height of the plant was estimated. A good correlation was found between the measured and estimated height using both stereo vision and the ultrasonic sensor. For the stereo vision sensor, r^2 between the maximum plant height and estimated height was 0.76. For the ultrasonic sensor, r^2 between the 25^{th} percentile of the group height statistics and plant collar height was 0.75.

Keywords. Stereo vision, corn, ultrasonic, precision agriculture, plant height

Introduction

Precision agriculture is a systems approach to manage spatial and temporal variability in crop fields for improved crop performance and environmental quality. Crop yield is affected by many factors such as soil properties, water availability, climatic variation, and topographic features. Since these factors interact with each other, it is difficult to identify their effects on crop growth and yield to determine how these factors could be managed site-specifically. A major limitation to identifying and mapping yield-limiting factors in fields is the availability of appropriate on-the-go sensing technologies for plant growth (Sadler et al.,

---

[1] Graduate Student
[2] Assistant Professor
[3] Professor

2000). Combine yield monitors are one example of sensing technology that has proven successful in measuring crop spatial variability in real-time with high spatial measurement density. The capability to measure crop growth rate in response to spatially varying factors at several times during the growing season would provide the opportunity to identify, map, and manage crop stress to minimize its effect on yield.

Sammis et al. (1988) used repeated, manual measurements of plant height as an indicator of water stress, evapotranspiration, and yield of irrigated corn, but they did not attempt to determine spatial patterns. Measuring spatial and temporal patterns of crop growth on a field scale has potential as a diagnostic tool for identifying crop stress or crop responses to spatial variability. However, traditional manual crop growth measurements are normally too labor intensive for sequential crop growth measurements to be used on a field scale for either commercial or research applications.

Image-based crop growth measurement has been shown to be effective in measuring and modeling crop plant growth in laboratory and greenhouse applications (Morden et al., 1997; Van Henten and Bontsema, 1995; Tarbell and Reid, 1991 a,b; Tarbell et al., 1991). However, vehicle-based sensing systems have not been developed to make repeated, non-destructive, non-contact crop growth measurements at field scales.

Tebourbi et al. (1999) developed a stereo vision system for measurement of soil texture and recommended that stereo vision be applied to crop growth sensing. McDonald et al. (1999) estimated soil surface roughness using stereo vision for different soil structures. They found good agreement between results obtained with the stereo vision approach and those obtained with other standard methods that were extremely time-consuming. Stereo vision has been applied to robotic harvesting applications to estimate the distance between the fruit or vegetable and the robotic manipulator (Takahashi et al., 1998; Kondo et al., 1996). Stereo vision has also been used in light interception studies of several crops (Ivanov et al.,1994; Sinoquet et al., 1998).

Kanuma et al. (1998) showed that a cabbage growth analysis system which employed stereo vision had improved measurement accuracy of leaf area over a single vision system. Matsuura et al. (2001) developed a transplant population growth analysis system that

estimated average height, leaf area, projected leaf area, and mass volume with good correlation to destructive measurements.

Ultrasonic sensing is another technology that has potential to sense corn height. Ultrasonic sensors have been used for measuring combine swath width (Sudduth et al, 1998; Vansichen et al., 1992). Ultrasonic sensing has also been investigated for measurement of crop biomass (Arca et al., 2000), canopies (Shibayama et al., 1985), and morphological characteristics (Ruixia et al., 1989).

The objectives of this research were to investigate and compare two sensing technologies, stereo vision and ultrasonic, as candidates for vehicle-based corn height sensors.

Methodology

Equipment

A lab-based sensor platform was developed which consisted of a motion control system on which a Pulnix TMC-9700 (Sunnyvale, CA) camera was mounted 2.1 m (6.9 ft) above and perpendicular to an imaging stage. A 6mm Tamron (Wayne, NJ) auto-iris (F1.6 to F360), varifocal (8-16mm) lens was used to focus on a 1.14 m (45 in.) by 0.86 m (33.75 in.) field of view (FOV). RGB component video from the camera was routed to a FlashBus MV color frame grabber board (Integral Technologies, Indianapolis, IN), which was installed in an Optiplex GX300 portable computer (Dell Computer Corp., Round Rock, TX) with dual 866 MHz Pentium III CPUs. The frame grabber had a resolution of 640 by 480 pixels and converted the analog video signal to 24-bit digital color images. Matlab version 6.1 (The Mathworks, Natick, MA) was used for image processing and plant height estimation.

The camera was moved in the horizontal plane using a lead screw in the y direction and a traction roller drive in the x direction (Figure 1). Camera motion was controlled with a PK 2110 (Z-World, Davis, CA) single-board computer with a 6.144 MHz microcontroller.

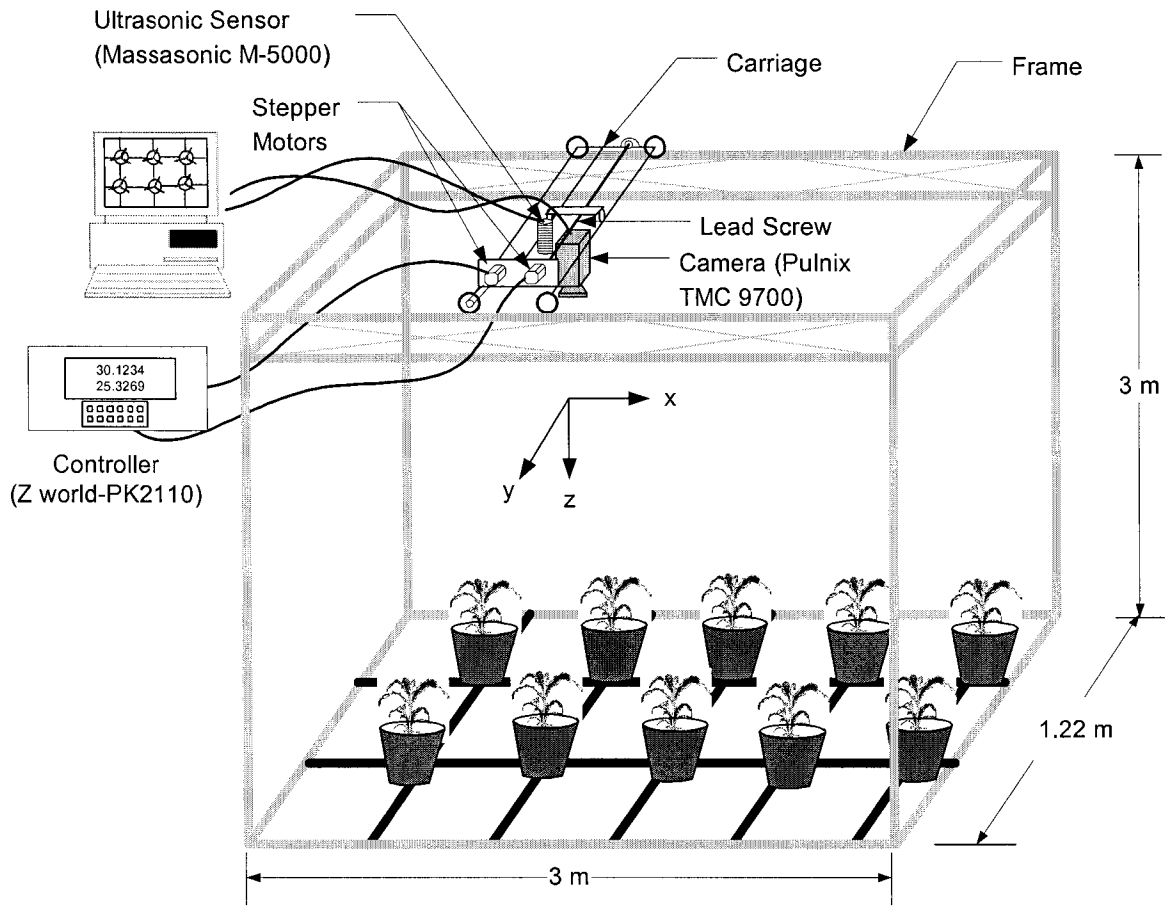


Figure 1. Experimental apparatus for camera calibration and depth measurements. The controller could position the camera with 1.59 mm precision in both x and y directions.

An ultrasonic sensor (Model M-5000/95, Massa Products, Hingham, MA) was mounted next to the camera on the sensing platform. This sensor operated at an ultrasonic frequency of 95 kHz and had an 8-degree conical beam angle and was designed to sense depth in a range between 0.3 to 4.0 m. A microprocessor was embedded in the sensor allowing settings such as sample rate to be programmed over an RS485 port. Sensor status such as temperature, depth to target, and signal strength could be read from the RS485 port. After poor preliminary results with the preprocessed depth to target measurement as an estimate of plant height, the analog echo detect monitor output from the sensor was used to estimate height. This output provided the peak detection signal of the ultrasonic signals including the transmitted ultrasonic burst as well as the received reflected signals from the targets in the field of view.

Procedure

Stereo vision algorithms typically consist of three steps: a) camera calibration – relating image pixels to a collection of points in the 3-D scene, b) point correspondence – determining the pair of points in the two images which relate to a single point in the scene, and c) 3-D reconstruction – finding the 3-D location of points in the scene (Sonka et al., 1999). The details of these steps are explained in this section.

Camera Calibration

Camera calibration is necessary to estimate the camera parameters. When a camera is calibrated, the intrinsic and extrinsic parameters of a camera are determined and used in succeeding calculations of height determination. If the coordinates of a point were x_{c1}, y_{c1}, z_{c1} with respect to the focal point of camera, the projection of the point on the CCD surface (U_1, V_1) is given by the equation:

$$\begin{bmatrix} U_1 \\ V_1 \\ 1 \end{bmatrix} = \begin{bmatrix} -f\,a & -f\,b & -u_0 \\ 0 & -f\,c & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \frac{x_{c1}}{z_{c1}} \\ \frac{y_{c1}}{z_{c1}} \\ 1 \end{bmatrix} = K \begin{bmatrix} \frac{x_{c1}}{z_{c1}} & \frac{y_{c1}}{z_{c1}} & 1 \end{bmatrix}^T \quad (1)$$

where the matrix K consists of the intrinsic parameters of the camera. Parameters a and c are scaling factors in pixel/mm and parameter b is the shearing component. Parameter f is the focal length of the camera and u_0 and v_0 are eccentricity parameters (Sonka, 1999). To implement a stereo imaging system, the camera being used must first be calibrated to correct for lens distortions and to relate pixel coordinates to corresponding unique rays in image space. The intrinsic camera parameters for the camera used in this research were determined using camera calibration software (Bouquet, 2002) based on the work of Zhang (1999), Heikkilä and Silven (1997), and Tsai (1987). Radial and tangential distortion coefficients of the image were also estimated through calibration.

Point Correspondence

Several image-processing steps were used to determine point correspondence of plants. First, image segmentation between plant and background was performed using the truncated ellipsoidal surface method developed by Shrestha and Steward (2001). This method

accomplished segmentation by using an ellipsoidal surface in RGB color space as a discrimination boundary between vegetation and non-vegetation regions.

Segmentation noise was removed through morphological opening with a 4-by-4 square structuring element. After this operation, objects segmented as plants larger than 100 pixels were labeled as plant objects. Next, the chain code for each valid plant object boundary was determined. Points on plant object edges were selected as possible candidate points for correspondence because of the large pixel intensity variance associated with the region around those points. Point correspondence was determined in two steps. In the first step, an initial estimation of object disparity was found by matching object boundaries in the two images. To determine the object disparity, pixel disparity error (PDE) was calculated. To calculate PDE, the boundary of the object was determined in the first image. Based on this boundary, the boundary of the search area in the second image was determined. Because the epipolar line was along the image columns, a search region was selected in the second image that was equal in width to the boundary of the object in the first image (Fig. 2). The object boundary pixel coordinates in each column of were compared with each boundary pixel in the same column of the second image.

Let us assume that for column i, mi numbers of boundary pixels were found at row number $y_{i11}, y_{i12},\ldots\ldots,y_{i1mi}$. For the same column i, ni numbers of boundary pixels were found in the search region at row $y_{i21}, y_{i22},\ldots.,y_{2ni}$. If disparity between two objects were s, then the PDE was calculated by summing the squared distance between every pixel in the same column of the object boundary to each pixel in the same column in the search boundary. Therefore, PDE for column i would be given by,

$$PDE_i = (y_{i11} + s - y_{i21})^2 + .. + (y_{i11} + s - y_{i2ni})^2 + .. + (y_{i1mi} + s - y_{i21})^2 + \ldots (y_{i1mi} + s - y_{i2ni})^2 \quad (2)$$

Total error was calculated by summing the PDE of the all columns 1 to k. The disparity s, which minimizes the total PDE, was calculated by differentiating the error with respect to s and equating to zero, which after simplification can be written as,

$$s = \frac{\sum_{i=1}^{k}\left\{mi\left(\sum_{j=1}^{ni} y_{i2j}\right) - ni\left(\sum_{j=1}^{mi} y_{i1j}\right)\right\}}{\sum_{i=1}^{k} mi \times ni} \quad (3)$$

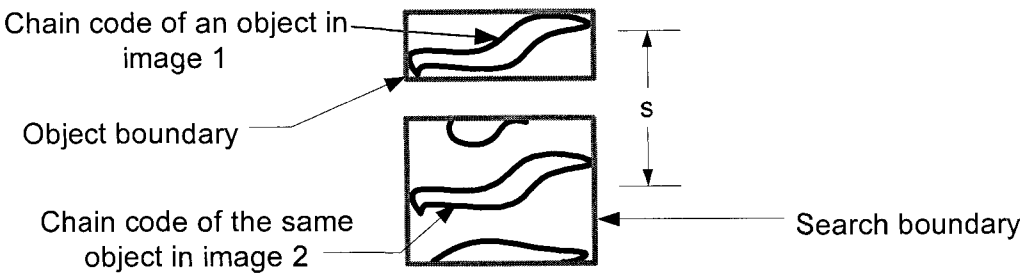


Figure 2. Chain code boundary matching of an object in an image pair. The separation distance d was calculated based on minimum error criteria.

Once the initial disparity estimate d was calculated, it was used to guide local matching of individual pixels. To find the disparity of the small section (ΔP) along the chain code, the object chain codes were divided into segments 10 column wide. For each selected sub-section of the chain code, a bounding rectangle was determined. It was ensured that there were at least 20 rows in the rectangle. Normalized cross-correlation of excess green (2×green – red – blue) images was used to find correspondence in the vicinity of the initial match. This not only sped up the searching process, but also led to point matching robustness. Robust correlation was insured by calculating the kurtosis and rejecting the match if kurtosis was less than a threshold of –1.

Height Estimation

Once the intrinsic camera parameters were known after calibration, depth information was calculated using the disparity of image coordinates taken from the camera located at two different viewpoints. It was assumed that the two camera viewpoints were on the same horizontal plane and separated only in the y direction by baseline distance, d. It was also assumed that the image coordinates for camera 1 and 2 were parallel, that is, there was no rotation between camera positions 1 and 2. Knowing the camera intrinsic parameters and disparity, the height of an object could be calculated using:

$$h = H - \frac{(fc)\,d}{\Delta P} \tag{4}$$

where H is the vertical distance from camera to a reference plane and h is the height of an object being measured from the reference plane. ΔP is the disparity in pixels of a control point in the two images.

Ultrasonic Sensing

The peak detection signal from the ultrasonic sensor was acquired at a sampling rate of 100 kHz by a data acquisition system (DaqBook/120, IOTech, Cleveland, OH). The sensor was mounted at a height of 2.1 m and was passed over individual plants at a velocity of 0.57 cm/s. The sensor was set for a measurement frequency of 2 Hz. The data acquisition took 1400 samples of the peak detection signal voltage for each scan. The number of scans for individual plants ranged from 252 to 360 depending on the projected area of the plant leaves.

Ultrasonic signal processing was done with MatLab script language. Signals were processed to find the start of echoes that were reflected back from the object in the FOV of the sensor. The time-of-flight of each detected signal was converted into a height estimate. The measurements associated with echoes from the pots and the imaging platform surfaces were filtered out. Height measurements from ten scans, corresponding to a 2.85 cm sensor shift, were grouped together and descriptive statistics such as mean, median, max, min, and 25^{th} percentile were calculated for each ten scan group. Groups that had less than 11 height measurements were filtered out to eliminate measurements taken at the extremes of the plant canopy.

Experimental Design

Sixteen corn plants at V3-V4 growth stage were placed on the imaging platform, and the camera was moved in the y direction at 5.1 cm increments with an image acquired at each point. Ten images were acquired. Then the camera was moved 37.5 cm in the x direction and another set of 10 images was acquired at 5.1 cm intervals. For a second set of 10 plants at V6, images were taken of individual plants at 2 inches intervals. Six images were taken of each plant, and multiple pairs of images were used as a replication for height estimation. The field of view (FOV) was 114 cm in the x direction resulting in an overlap of 75 cm from one camera location to the next. The maximum height of each corn plant was measured manually to compare the results with calculated height. Plant height ranged from 35 cm to 100 cm including the pot height.

For ultrasonic sensing, 2 sets of 10 corn plants, at V6 and V9 growth stages respectively, were placed on the imaging stage one plant at a time and scanned with the ultrasonic sensor. Temperature at scanning was recorded and ranged between 27° and 29°C.

Manual measurements were taken of corn plant collar heights, maximum plant heights, and maximum height of each leaf. Each leaf was classified according to its orientation, and digital still photos were taken of each plant for future reference.

Results and Discussion

Stereo Vision Sensing

Camera calibration revealed that there was both radial and tangential distortion in the images. The K matrix in equation (3) was determined. The value of f·c, the camera focal length times the scale factor in equation 4, was found to be 1163.86 ± 5 pixels.

The image had pincushion distortion. A rectified image is shown in Figure 3. The barrel-shaped image outline is due to distortion correction. The height was calculated with the maximum disparity detected for an object. When there were more than one segmented object belonging to the same plant, the height of the highest object was used as the height for that plant. For instance, in Figure 3, 3 objects were detected for 2 plants. A height of 52.7 cm was assigned for the plant in the top row and 65.8 cm for the plant in the middle of the bottom row.

Regression analysis was used to describe the relationship between manually measured plant height and optically measured height (Figure 4). The r^2 value of the linear regression line was 0.76. The slope and intercept of the regression line with 95% CI's were 1.04 ± 0.08 and -12.41 ± 5.6, respectively. The RMSE of the model was found to be 10.56 cm. The cluster in the lower region corresponded to the plants at V3-V4 stage. The data points were more scattered for measurements taken at the V3-V4 stages because during that part of the experiment height was manually measured as the maximum height from ground to the highest point on the plant. At these growth stages the highest point on the plant is often an erect leaf or leaf in the whorl that is not fully extended. At the V6 stage, height was manually measured as the maximum height of the uppermost fully extended leaf with a planophile or horizontal orientation. It was found that the optically measured height was more correlated to the latter measurement. In addition, the optically measured average leaf height was poorly correlated to the manually measured collar height. Part of the reason for this poor correlation was the small range of available manual collar measurements.

For height estimation, exact pixel matching is very critical. However during stereo matching, the disparity measurement ΔP may vary slightly because of several error sources, such as image distortion, camera calibration error, and camera position misalignment. Even under ideal conditions, the minimum disparity resolution results in a corresponding minimum height estimation resolution. The height measurement resolution under the assumption of one pixel mismatch can be calculated by differentiating equation 4, which after simplification reduces to,

$$\frac{dh}{d\Delta P} = \frac{(H-h)^2}{(fc)d} \tag{5}$$

Thus, the height measurement resolution for a given camera is directly proportional to the square of the distance of an object from the camera and inversely proportional to the camera separation. For this reason, images one image apart in the image series taken at 5.1 cm apart were used as stereo pairs with a 10.2 cm baseline distance for height estimation. For height measurements of 35 to 100 cm in this experiment, the minimum height resolution due to one pixel matching resolution was about 3 cm.

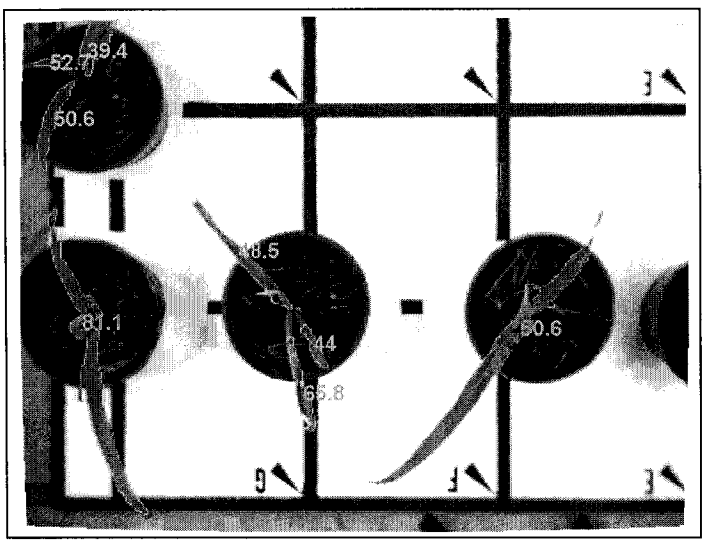


Figure 3. Undistorted image showing chain coding after plant detection and height estimation. The numerical values indicate the estimated maximum height of each object.

Ultrasonic Sensing

The mean of height measurements using the ultrasonic sensing system tended to be less correlated with the manual maximum height measurements than those taken with the

optical system. Correlation analysis (Table 1) showed that the maximum, minimum, median, and 25th percentile of the ultrasonic measurements from the groups of 10 had correlation coefficients ranging from 0.60 to 0.66 with the maximum measured heights. The maximum measured heights were often taken from the tips of erect leaves, which were usually leaves that were not fully extended and were in the plant whorl. Erect leaves were not detected as clearly by the ultrasonic sensor as leaves that had a more horizontal orientation because these leaves reflected the ultrasonic signal away from the sensor (Figure 5).

The maximum height of the uppermost fully extended leaf with a horizontal orientation was slightly more correlated with ultrasonic group statistics than maximum height, with correlation coefficients ranging from 0.62 to 0.75. Because these leaves had an approximately horizontal surface at the top of the bend, they often produced a strong echo signal.

The ultrasonic measurements were well correlated with the measured collar height with correlation coefficients ranging from 0.77 to 0.85. When the 25th percentile of the group height statistics were regressed on the collar height measurement, the regression line had a slope of 0.97 and a y-intercept of 3.16 with $r^2 = 0.75$ (Fig. 6).

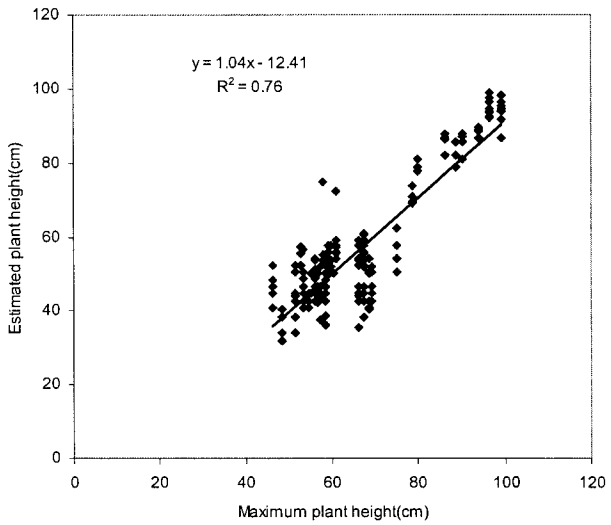
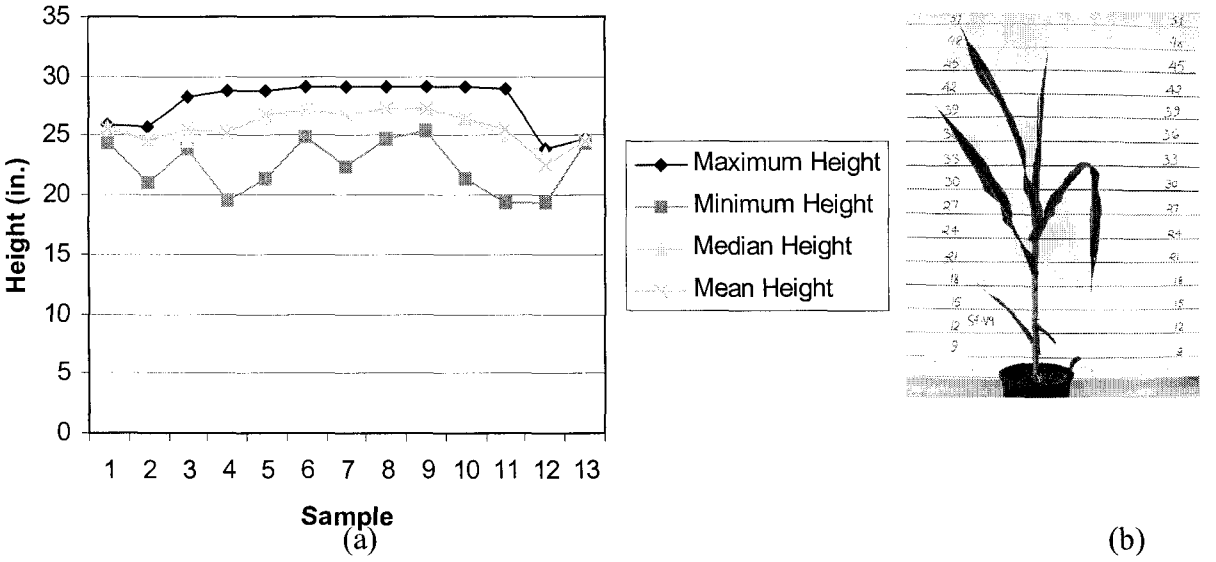


Figure 4. Estimated plant height from stereo vision system versus manually measured plant heights for V3-V6 growth stage corn plants. N = 197.

Table 1. Correlation coefficients of corn plant height measurements. N = 389.

	Max height	Min. height	Median height	Mean height	25th percentile height	Max. meas. height	Max. Bent meas. height	Meas. collar height
Max height	1							
Min height	0.74	1						
Median height	0.90	0.90	1					
Mean height	0.93	0.91	0.98	1				
25th percentile height	0.86	0.93	0.96	0.98	1			
Maximum measured height	0.63	0.60	0.65	0.66	0.66	1		
Max. bent meas. height	0.75	0.62	0.73	0.74	0.71	0.81	1	
Meas. collar height	0.77	0.80	0.83	0.84	0.85	0.87	0.78	1



(a) (b)

Figure 5. (a) Ten ultrasonic scan group statistics as a corn plant (b) was scanned from left to right. Note that the tall erect leaves are not appearing in the measurements.

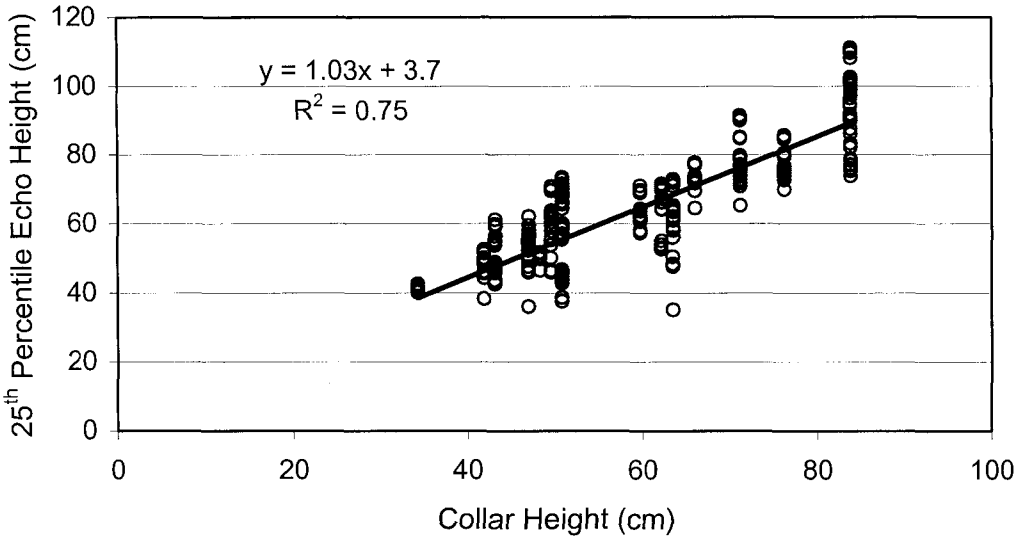


Figure 6. Twenty-fifth percentile of ultrasonic scan group statistics regressed onto manually measured collar height.

Conclusion

Two methods were developed and compared for corn plant height estimation. A single camera was used on an imaging platform to generate stereo image pairs. The movement of the camera was precisely controlled using a stepper motor controlled by a digital controller. Corn images from V3 to V6 stages were acquired with the camera with known camera movements.

An image correspondence algorithm was developed, and an initial estimate of object disparity was calculated minimizing the total PDE for chain code. After determining the initial disparity estimate, the disparity for chain code sub-sections was calculated using excess green images. The matching was further verified by calculating the kurtosis of the cross correlation values along the base line.

Optically measured plant heights using stereo vision were moderately correlated with manually measured maximum plant height. However, the optically measured heights tended to be lower because the erect leaves in the whorl could not be detected as separate objects in the segmentation process and hence the measured height was less than the highest point. Pixel matching resolution is a major source of uncertainty in the height measurements.

Subpixel interpolation for pixel matching will be investigated for future improvements in measurement accuracy.

Ultrasonic measurement was well correlated with collar height. The r^2 of the regression line between 25th percentile of group statistics and collar height was 0.75. When plants were placed too close, however it was difficult to separate echoes from individual plants.

This research suggests that the two sensing systems may be complementary in sensing corn plant height. In the future, it might be possible to use both ultrasonic sensing and stereo vision in combination to achieve improved plant height estimates.

Acknowledgements

Support for this project came from the Iowa State University Research Grant program. Thanks to Nicholas Krueger and Eric Stephan for their assistance on this research.

References

Arca, B., F. Benincasa, A. Materassi, and A. Ventura. 2000. Project and construction of an instrument for a non-destructive determination of biomass. *Rivista di Ingegneria Agraria* 31(4): 201-206.

Bouquet, J.Y. 2002. Camera calibration toolbox for MatLab. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html. Accessed on 12 Dec. 2001.

Heikkilä, J. and O. Silven. 1997. A four-step camera calibration procedure with implicit image correction. In *Proceedings of 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 1106-1112. San Juan, Puerto Rico. 17-19 June.

Ivanov N., P. Boissard, M. Chapron, P. Valery. 1994. Estimation of the height and angles of orientation of the upper leaves in the maize canopy using stereovision. *Agronomie* 14(3): 183-194.

Kanuma, T., K. Ganno, S. Hayashi, and O. Sakaue. 1998. Leaf area measurement using stereo vision. *Proc. of the 3rd IFAC/CIGR workshop on Artificial Intelligence in Agriculture, Makuhari Messe, Chiba, Japan.* 157-162. 24-26 April.

Kondo N., Y. Nishitsuji, P.P. Ling, and K. C. Ting. 1996. Visual feedback guided robotic cherry tomato harvesting. *Transactions of the ASAE* 39(6): 2331-2338.

Matsuura, Y., D. He, and T. Kozai. 2001. Development of a transplant population analysis system. *ASAE Paper No. 01-7051.* St. Joseph, MI: ASAE.

McDonald, A. J., S. Crossley, J. C. Bennett, G. Cookmartin, K. Morrison, and S. Quigan. 1999. Stereo vision measurements of soil surfaces and their utility in testing microwave backscattering models. IEEE *International Geoscience and Remote Sensing Symposium, IGARSS Proceedings*. Vol 5: 2634 -2636.

Morden, R. E., P.P. Ling, and G. A. Giacomelli. 1997. An automated plant growth monitoring systems using machine vision. *ASAE Paper No. 974033*. St. Joseph, MI: ASAE.

Ruixiu, S., J. B. Wilkerson, L.R. Wilhelm, and F. D. Tompkins. 1989. A microcomputer-based morphometer for bush-type plants. *Computer and Electronics in Agriculture* 4(1): 43-58.

Sadler, E.J., P.J. Bauer, W.J. Busscher. 2000. Site-specific analysis of a droughted corn crop: I. Growth and grain yield. *Agronomy Journal*. 92:395-402.

Sammis, T.W., D. Smeal, and S. Williams. 1988. Predicting corn yield under limited irrigation using plant height. *Transactions of the ASAE* 31:830-838.

Shibayama, M., T. Akiyama, and K. Munakata. 1985. A portable field ultrasonic sensor for crop canopy characterization. *Remote Sensing of Environment* 18(3): 269-279.

Shrestha, D.S and B.L. Steward 2001. Automatic corn plant population measurement using machine vision. *ASAE Paper No. 01-1067*. St. Joseph, MI: ASAE.

Sinoquet, H., S., Thanisawanyangkura, H. Mabrouk, and P. Kasemsap. 1998. Characterization of the light environment in canopies using 3D digitising and image processing. Annals of Botany 82(2): 203-212.

Sonka, M., V. Hlavac, and R. Boyle. 1999. *Image Processing, Analysis, and Machine Vision*. Pacific Grove, CA: Brooks/Cole Publishing Co.

Sudduth, K. A., S. T. Dummond, W.W. Wang, and M.J. Krumpelman. 1998. Ultrasonic and GPS measurement of combine swath width. *ASAE Paper No. 983096*. St. Joseph, MI.

Takahashi, T., S. Zhang, M. Sun, and H. Fukuchi. 1998. New method of image processing for distance measurement by a passive stereovision. *ASAE Paper No. 983031*. St. Joseph, MI: ASAE.

Tarbell, K. A. and J. F. Reid. 1991a. A computer vision system for characterizing corn growth and development. *Transactions of the ASAE* 34(5): 2245-2255.

Tarbell, K. A. and J. F. Reid. 1991b. Spatial and spectral characteristics of corn leaves collected using computer vision. *Transactions of the ASAE* 34(5): 2256-2263.

Tarbell, K. A., D. K. Tcheng, and J. F. Reid. 1991. Corn growth and development attributes obtained using inductive learning techniques. *Transactions of the ASAE* 34(5): 2264-2271.

Tebourbi, R., Z. Belhadj, M. Zribi, and M. R. Boussema. 1999. 3D reconstruction of natural targets by stereovision. *IEEE 1999 International Geoscience and Remote Sensing* Symposium, IGARSS '99 Proceedings. Vol 2: 1128 –1130.

Tsai, R. Y. 1987. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation* 4(3): 323-344.

Van Henten, E. J. and J. Bontsema. 1995. Non-destructive crop measurement by image processing for crop growth control. *Journal of Agricultural Engineering Research* 61: 97-105.

Vansichen, R. and J. DeBaerdemaeker. 1992. Measuring the actual cutting width of a combine by means of an ultrasonic distance sensor. *Proc. of Trends in Agric. Eng.* Conference. 615-621 Prague. 15-18 Sept.

Zhang, Z. 1999. Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In *Proc. of the 7th IEEE International Conference on Computer Vision*. 666-673. Kerkyra, Greek. 20-27 Sept.

General Conclusions

This research provided the basis of machine vision system for early stage corn plant location and plant height measurement. The uncontrolled environment of agricultural fields posed many challenges in developing a standard procedure that works with all different field conditions. Natural lighting variation in the field poses a challenge in developing a plant segmentation algorithm. Modern cameras are capable of adjusting the camera aperture to a wide range of lighting conditions; however, in extreme lighting conditions, this automatic aperture may not be sufficient.

One of the objectives of this research was to develop a robust plant segmentation system which could separate plant from the background in various lighting condition. The truncated ellipsoidal (TE) decision surface (Discussed in chapter 3) segmented the plant from background in a wide range of lighting conditions generally encountered in the field. However, when the sky was too dark or too bright, the segmentation performance degraded. An artificial neural network was able to adjust the segmentation parameters to different lighting condition for a better segmentation but, again in extreme lighting variation, the error rate of segmentation was high. Whether it was an optimum parameter or constant parameter analysis, TE decision surface always had a lower rate of segmentation error than NDI or Bayesian methods of segmentation.

Second objective of this research was to develop a machine vision algorithm to locate and count the individual plants from the segmented image. An iterative algorithm was developed with manually optimized parameters to detect, count and measure inter-plant spacing. The algorithm performed well in relatively small amount of data set, however, error increased when adapted to different operational conditions with higher weeds and plant sizes. Therefore the algorithm was refined to exclude weeds and calculate the parameters value automatically using first and second order statistics of the extracted features. The modified algorithm performed well in a wide range of operating conditions. However, the performance of the plant counting algorithm was sensitive to the video quality. Some segments of the video could not be processed because of poor video quality. Video quality was a composite result of many factors like lighting condition, camera aperture adjustment, digital gain, and

vibration. Dirty video head was also responsible for degrading video quality in dusty environment.

At higher vehicle speeds, it was observed that the video frames were motion blurred. At higher vibration, the video sequencing algorithm not only took longer time to find a statistically significant match but also a frequent erroneous matching of the video frames was observed. Since the camera mounting was an un-damped system, excessive vibration at higher vehicle speed was observed. High vibration made the image blurred and sequencing was more erroneous.

The third objective was to develop an image sequence mosaicing procedure. A statistically robust image correspondence algorithm was developed based on a patch-matching algorithm. Because of the iterative nature of image correspondence algorithm, the time required to match two subsequent images varied. However, it was found that ESCOPE was able to sequence the images while meeting real time constraints. The image correspondence technique was accurate enough to be used to measure the interplant distances by counting pixels between two plants. Both pixel counting and GPS interpolation could be used for plant location and interplant spacing measurement. Pixel counting was used for plant spacing measurements and the results were compared with manual measurement. The GPS was not used for this purpose because RTK GPS signal was missing from some sections.

Following the development of the plant detection and spacing measurement algorithm, a general component based software architecture was developed so that the data acquisition and processing system could be extended to incorporate other field variabilities. This met the fourth objective of this research. Following the architecture, ESCOPE software was developed. ESCOPE could communicate with and control a digital camcorder through IEEE1394 port. ESCOPE could be used to count the corn plants and measure the interplant distances. ESCOPE output the plant variables such as corn plant location, projected canopy area and interplant distances in a text file. The outputted text file can be imported into GIS software to generate a field variability map.

The fifth objective was to develop a robust correspondence method for height determination. Plant height determination requires the images taken from two known camera to a common scene of interest. Feature correspondence was the key problem in stereo vision.

A chain code based unique feature correspondence technique was developed and tested to measure the plant height. We have used a single camera and displaced the camera by a known distance to take another image. Even though the movement was precisely controlled using a lead screw, we observed some eccentricity in screw rotation. This wobbled the camera and introduced some error in height measurement. Plant height measured using stereovision technique was moderately correlated with manually measured maximum plant height. It was also revealed that pixel matching resolution was a major source of uncertainty in the height measurement. Each of the developed algorithms was tested and the results were compared with manually measured values. The results are described in the relevant sections of this paper chapters. This met the last objective of this research work.

Future Recommendations:

This research has a great potential to be expanded to make the precision agriculture field application feasible and more profitable. More field variables can be added into the software architecture and measure and study the effect of there field variables on final yield. However based on the experience I have gained from this research work, following work has been recommended:

1. Develop a camera mounting system to minimize the vibration.
2. Use a translucent cover during data acquisition to diffuse the light, to avoid shadow.
3. Improve on image correspondence speed to use in real time processing.
4. Make the ESCOPE more robust by taking account of more exceptional cases.
5. Test stereovision technique to determine plant height in the field using two separate cameras.
6. Incorporate stereo vision techniques into the ESCOPE.

References

Andreasen C., M. Rudemo, and S. Sevestre 1997. Assessment of weed density at an early stage by use of image processing. *Weed Research* 37(1): 5-18.

Berck, P. and G. Helfand. 1990. Reconciling the Von Liebig and differentiable crop production functions. *American Journal of Agricultural Economics* 72(4): 985-996.

Berck, P., J. Geoghegan, and S. Stohs. 2000. A strong test of Von Liebig hypothesis. *American Journal of Agricultural Economics* 82(4):948-955.

Colvin, T.S., S. Arslan, and D.W. Meek. 1999. Yield monitors, combines, and their interactions. *Society of Automotive Engineers Technical Paper Series*, Paper No. 1999-01-2846, Warrendale, PA.

Duncan, W.G. 1958. The relationship between corn populations and yield. *Agronomy Journal* 50:82-84.

El-Faki, M. S., N. Zhang, and D. E. Peterson. 2000. Weed detection using color machine vision. *Transactions of the ASAE* 43(6): 1969-1978.

Gonzalez, R.C. and R.E. Woods, 1992, Digital Image Processing, Addison-Wesley Publishing Company, Reading, MA.

Hetzroni, A. and G. E. Miles. 1992. Machine vision monitoring of plant health. In International Winter Meeting. (December 1992): *American Society of Agricultural Engineers*.

Ivanov N., P. Boissard, M. Chapron, P. Valery. 1994. Estimation of the height and angles of orientation of the upper leaves in the maize canopy using stereovision. *Agronomie* 14(3): 183-194.

Jia, J., G. W. Krutz and H.G. Gibson. 1990. Corn plant locating by image processing. Proc. SPIE, *Optics in Agriculture* vol. 1379: 246-253.

Kanuma, T., K. Ganno, S. Hayashi, and O. Sakaue. 1998. Leaf area measurement using stereo vision. *Proc. of the 3rd IFAC/CIGR workshop on Artificial Intelligence in Agriculture, Makuhari Messe, Chiba, Japan.* 157-162. 24-26 April.

Kayode, G. O. and A. A. Agboola. 1981. Effect of different nitrogen levels, plant population and soil nutrient status on yield and yield components of maize (Zea mays L.) in different ecological zones of Nigeria. *Fertilizer Research* 2(3): 177-191.

Kondo N., Y. Nishitsuji, P.P. Ling, and K. C. Ting. 1996. Visual feedback guided robotic cherry tomato harvesting. *Transactions of ASAE* 39(6): 2331-2338.

Lines, J.A., R.D. Tillett, L.G. Ross, D. Chan, S. Hockaday, and N.J.B. McFarlane. 2001. An automatic image-based system for estimating the mass of free-swimming fish. *Computer and Electronics in Agriculture* 31(2): 151-168.

Ling, P.P., G.A. Giacomelli and T.P. Russell. 1995. Monitoring of plant development in controlled environment with machine vision. *Advances in Space Research*, 18(4-5):101-112.

Mathers, A. C. and B. A. Stewart. 1982. Sunflower nutrient uptake, growth, and yield as affected by nitrogen or manure, and plant population helianthus annuus. *Agronomy Journal* 74(5): 911-915.

Matsuura, Y., D. He, and T. Kozai. 2001. Development of a transplant population analysis system. *ASAE Paper No. 01-7051.* St. Joseph, MI: ASAE.

McDonald, A. J., S. Crossley, J. C. Bennett, G. Cookmartin, K. Morrison, and S. Quigan. 1999. Stereo vision measurements of soil surfaces and their utility in testing microwave backscattering models. *IEEE International Geoscience and Remote Sensing Symposium, IGARSS Proceedings*. 5: 2634 -2636.

Meyer, G. E., T. Mehta, M. F. Kocher, D. A. Mortensen, and A. Samal. 1998. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. *Transactions of ASAE* 41(4):1189-1197.

Morden, R. E., P.P. Ling, and G. A. Giacomelli. 1997. An automated plant growth monitoring systems using machine vision. *ASAE Paper No. 974033*. St. Joseph, MI: ASAE.

National Research Council. 1998. *Precision Agriculture in the 21st Century: Geospatial and Information Technologies in Crop Management*. National Academic press: Washington D.C.

Nichols, S.W. 2000. Method and apparatus for counting crops sensor. U.S. Patent No. 6,073,427.

Nielsen, R.L. 1995. Planting seed effects on stand establishment and grain yield of corn. Journal of Production Agriculture. 8(3): 391-393.

Nishiwaki, K., T. Togashi, K. Amaha, and K. Matsuo. 2001. Estimate crop position using template matching in rice production. *ASAE Paper no. 013103*. St. Joseph, Mich.: ASAE.

Otsu, N. 1979. A threshold selection method from gray - level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9(1): 62-66.

Pérez A.J., F. López, J. V. Benlloch, and S. Christensen. 2000. Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture* 25(3): 197-212

Rees, G., 1999, The Remote Sensing Data Book, Cambridge University Press, Cambridge, UK, pp. 245-6.

Robert, P. C. 2000. Site specific management for the twenty-first century. *Horticulture* 10(3): 444-447.

Sadler, E.J., P.J. Bauer, W.J. Busscher. 2000. Site-specific analysis of a droughted corn crop: I. Growth and grain yield. *Agron J.* 92:395-402.

Sammis, T.W., D. Smeal, and S. Williams. 1988. Predicting corn yield under limited irrigation using plant height. *Transactions of the ASAE* 31:830-838.

Sanchiz, J.M., F. Pla, J.A. Marchant and R. S. Brivot. 1995. Structure from motion techniques applied to crop field mapping. *Image and Vision Computing* 14(5): 353-363.

Shrestha, D.S and B.L. Steward 2001. Automatic Corn Plant Population Measurement Using Machine Vision. *ASAE Paper No. 01-1067.* St. Joseph, MI: ASAE.

Sinclair, T. R. and W. I. Park. 1993. Inadequacy of the Liebig limiting paradigm for explaining varying crop yield. *Agronomy Journal* 85(3): 742-746.

Sinoquet, H., S., Thanisawanyangkura, H. Mabrouk, and P. Kasemsap. 1998. Characterization of the light environment in canopies using 3D digitising and image processing. *Annals of Botany* 82(2): 203-212.

Sonka, M., V. Hlavac, and R. Boyle. 1999. *Image Processing, Analysis, and Machine Vision.* Pacific Grove, CA: Brooks/Cole Publishing Co.

Steward, B. L. and L. F. Tian. 1998. Real-time weed detection in outdoor field conditions. *In Proc. SPIE 3543, Precision Agriculture and Biological Quality,* eds. G. E. Meyer and J. A DeShazer, 266-278. Bellingham, Wash.: SPIE.

Steward, B. L., L. F. Tian, and L. Tang. 1999. Detection of outdoor lighting variability for machine vision-based precision agriculture. *ASAE Paper No. 99-3032*, St. Joseph, MI.

Strock, J. S. and G. L. Malzer. 2000. Spatial variability of corn yield affected by landscape position and mechanical impedance. In *Proceedings of fifth international conference on precision agriculture* P.C.Robert, R.H.Rust, and W.E.Larson, eds. Bloomington, MN, USA.

Sudduth, K.A., S.J. Birrell, M.J. Krumpelman. 2000. Field evaluation of a corn population. Proceedings *of the Fifth International Conference on Precision Agriculture.* Eds. T.C. Robert, R.H. Rust, W.E. Larson. July 16-19, Bloomington, Minnesota USA.

Takahashi, T., S. Zhang, M. Sun, and H. Fukuchi. 1998. New method of image processing for distance measurement by a passive stereovision. *ASAE Paper No. 983031.* St. Joseph, MI: ASAE.

Tang, L., L. Tian, and B. L. Steward. 2000. Color image segmentation with genetic algorithm for in field weed sensing. *Transactions of the ASAE* 43(4): 1019-1027.

Tang, L., L. Tian (a). 2002. Machine vision for automated corn plant spacing, growth stage and population measurements - part I: real-time image-sequencing. *ASAE Paper No.023099.* St. Joseph, MI: ASAE.

Tang, L., L. Tian (b). 2002. Machine vision for automated corn plant spacing, growth stage and population measurements - part II: plant identification. *ASAE Paper No.023100.* St. Joseph, MI: ASAE.

Tarbell, K. A. and J. F. Reid. 1991a. A computer vision system for characterizing corn growth and development. *Transactions of the ASAE* 34(5): 2245-2255.

Tarbell, K. A. and J. F. Reid. 1991b. Spatial and spectral characteristics of corn leaves collected using computer vision. *Transactions of the ASAE* 34(5): 2256-2263.

Tarbell, K. A., D. K. Tcheng, and J. F. Reid. 1991. Corn growth and development attributes obtained using inductive learning techniques. *Transactions of the ASAE* 34(5): 2264-2271.

Tebourbi, R., Z. Belhadj, M. Zribi, and M. R. Boussema. 1999. 3D reconstruction of natural targets by stereovision. *IEEE 1999 International Geoscience and Remote Sensing Symposium, IGARSS '99 Proceedings.* Vol 2: 1128 –1130.

Tian, L. F. and D. C. Slaughter. 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation.*Comp. & Elect. in Agric.*21(3):153-168.

Van Henten, E. J. and J. Bontsema. 1995. Non-destructive crop measurement by image processing for crop growth control. *Journal of Agricultural Engineering Research* 61(2): 97-105.

Ward, B. D. and M. S. Cox. 2000. Influences of soil chemical and physical properties on site-specific cotton production. In *Proceedings of fifth international conference on precision agriculture* P.C.Robert, R.H.Rust, and W.E.Larson, eds. Bloomington, MN, USA.

Wiley, R.W and S.B. Heath. 1970. The quantitative relationship between plant population and crop yield. *Advances in Agronomy* 21:281-321.

Will, J. 2001. Sensor Fusion for field robot localization. Ph.D. diss. *University of Illinois at Urbana-Champaign*, Champaign.

APPENDIX A

Camera Basics

Lens

In machine vision calculations it is customary to consider a camera as a pinhole model. A pinhole model is essentially the same as using a thin lens camera. An ideal thin lens has following characteristics:

1 Parallel rays of light passing through a double convex lens meet at the focal point of or focus designated by F.

2 Any ray passing from the optical center of the lens passes un-deflected through the lens.

Following these two simple rules we can derive the lens equation as follows. Let an object is placed at distance u in front of lens at point A. The tip of the object B emits light rays in directions but in order to make a sharp image, all the rays passing through lens must converge at point D. Among all combinations of rays, ray BC and ray BO is of special importance in determining the position of the image point D. BC is parallel to optical axis AE. From the lens characteristics, ray BC will pass through focal point F and ray BO will pass through lens un-deflected. The point at which the rays meet is the point where the image is formed. Let us assume that focal length of the lens is f and the distance from lens to image plane is v.

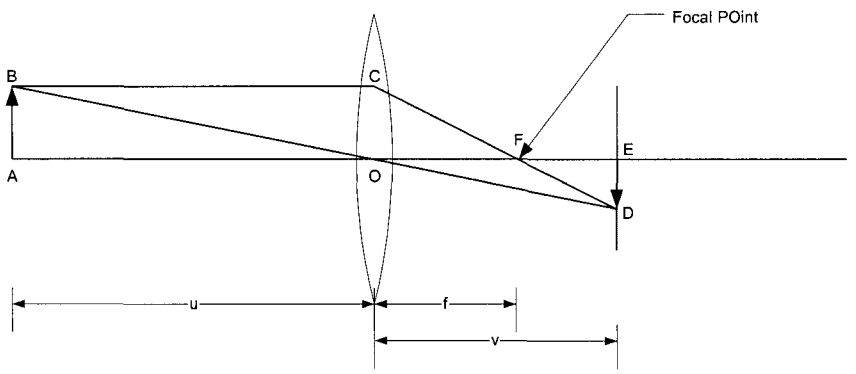


Figure A1. Schematics of a simple lens

From similar triangles, ABO and OED, we have, $\frac{AB}{ED} = \frac{u}{v}$ and from similar triangles, OCF and DEF, we have, $\frac{AB}{ED} = \frac{f}{v-f}$. From these two equations we can derive well known lens equation:

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \tag{A1}$$

F-Stop

In machine vision applications, image quality plays a vital role in image post processing. The F stop is a common measure of aperture opening that controls the amount of light falling into image plane. The F-stop is defined as:

$$F = \frac{f}{a} \tag{A2}$$

Where, F is the F number, f is focal length and a is the effective diameter of the lens. The effective diameter of the lens is the diameter of the iris as seen from in front of the lens. Each succeeding F- stop reduces the effective diameter of lens by half. The common F- stops are 1,1.4,2,2.8,4,5.6,8,11,16,22,32. Notice that larger the lens opening diameter, the lower the value of F-stop for a given lens. Cameras are designated by their minimum F-stop and focal length like: 16mm/1.4.

Depth of field

The lens creates an image of an object in the plane of the image screen (fig. A2). The lens has a focal length f, iris diameter $a\ (=f/F)$. As in the previous case, The image screen is located at the distance v from the lens and the object is at the distance u from the lens, on the other side from the image screen. Suppose the object image is in focus, then the distance between the lens and the image plane depends on the focal length and the distance to the subject given by equation (A1).

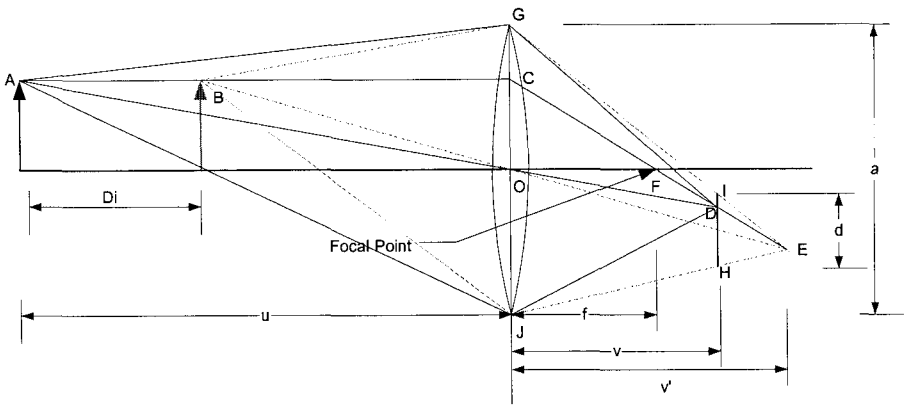


Figure A2. Depth of field when object is moved towards the lens.

Now imagine that the object point A is shifted inwards to point B along optical axis by *Di*. The point B would be focused at point E at v' distance from lens center. Because of this moving the object be now defocused and hence forms a circle of diameter d on image plane instead of a sharp point. From the lens equation we have,

$$\frac{1}{u - D_i} + \frac{1}{v'} = \frac{1}{f} \tag{A3}$$

Again, from similar triangle JGE and HIE we can derive following relationship. The ratio of JG and HI is equal to the ratio of JE and HE which in turn are equal to ratio v' and v'-v. Therefore,

$$\frac{a}{v'} = \frac{d}{v' - v} \tag{A4}$$

Substituting, $a = f/F$ and solving for v', which, after simplification using equation A3 and A1 gives the relationship:

$$D_i = \frac{udF(u - f)}{f^2 + dF(u - f)} \tag{A5}$$

In practice the allowable diameter of defocused circle is not equal to zero. The maximum allowable diameter is determined by hardware used and also depends on specific application. This maximum diameter is generally called the circle of least confusion. From above figure, it is clear that as object moves inward point E will move further from lens along the line CD. To keep the circle of least confusion within limit we need to decrease the iris diameter. Equations A5 also tell us that as for a given d, D_i increases as u increases in a non linear fashion.

Similarly, when point A is shifted outside by distance D_o, the point will focus in front of image plane and point B will form a circle on image plane of diameter d.

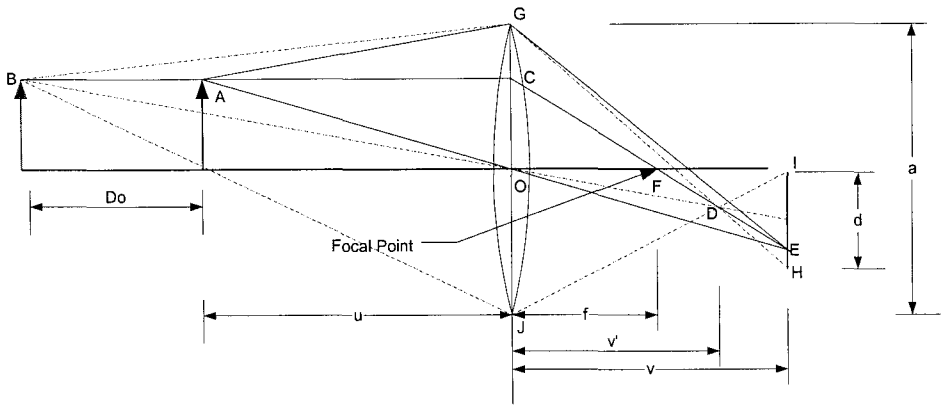


Figure A3. Depth of field when an object is moved away from the lens.

From lens equation:

$$\frac{1}{u+D_o}+\frac{1}{v'}=\frac{1}{f} \tag{A6}$$

From similar triangles, JGD and DIH,

$$\frac{a}{v'}=\frac{d}{v-v'} \tag{A7}$$

Solving equation A1, A2, A6 and A7 for D_o gives:

$$D_o=\frac{udF(u-f)}{f^2-dF(u-f)} \tag{A8}$$

The total length $D_i + D_o$ is called depth of field where the picture will be sharply focused in the image plane. The knowledge of the depth of field is important in determining the precision of the stereo vision system. In stereovision, different objects are at different distances from the camera and it is not possible to get the sharply focused image at all depths. The depth accuracy of a scene point degrades as the distance of the point from the point at focus increases. In order to minimize the effect, larger aperture number (or smaller aperture size) should be used. This requires the higher lighting intensity falling into the surface being imaged.

Appendix B

Stereo Vision and camera calibration

Recording images using a camera is equivalent to mapping object point X in the object space to image point I in the image plane. Two coordinate systems are defined (Fig B1) as world coordinate system (with subscript w) and camera coordinate system (with subscript c).

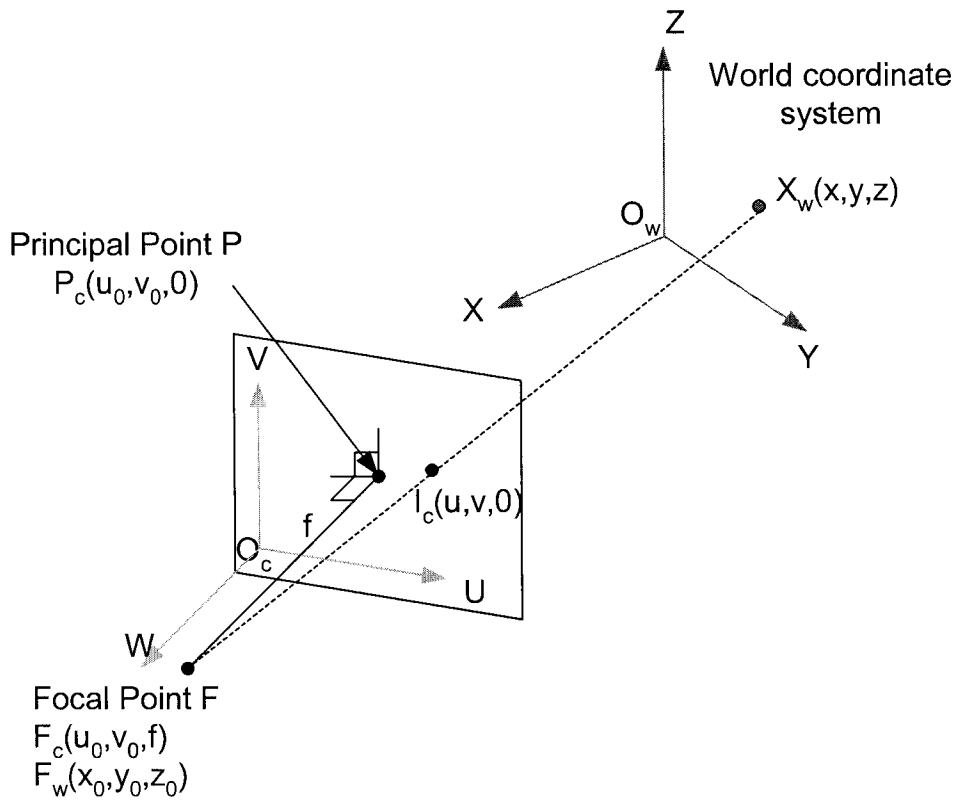


Figure B1. A world coordinate point X_w is projected into film plane I with camera coordinate $(u,v,0)$

Each of the points can be either represented in world coordinate or camera coordinate. For instance the focal point F has camera coordinates F_c and world coordinates F_w. The principal point is defined as a point intersecting the line perpendicular to image plane (parallel to W axis) and passing through the focal point and image plane. Assuming the camera coordinates of the principal point to be $[u_o, v_o, 0]$, the position of point F in the camera coordinate for point F becomes $[u_o, v_o, f]$. Vector **FI** drawn from point F to I is $[u-u_o, v-v_o, -f]$.

Since point F, I and X are collinear, **FI** = m **FX**, where m is a scaling factor. In order to directly relate the coordinates, it is necessary to describe them in a common reference frame. One good way to do this is to transform vector **FX** to the camera coordinate system. If R be the rotational matrix from world coordinate to camera coordinate system

$$\mathbf{FX}_c = R\,\mathbf{FX}_w \tag{B1}$$

$$\mathbf{FI}_c = m\,\mathrm{FX}_c = mR\,\mathbf{FX}_w \tag{B2}$$

$$\begin{bmatrix} u-u_0 \\ v-v_0 \\ -f \end{bmatrix} = m \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix} \tag{B3}$$

In actual case there will be some optical errors from the lens. The units of measurement for world coordinate system which may be in meters can be different than camera coordinate system which may be in pixels. In addition, in actual CCD, the aspect ratio of pixel in two directions may be other than 1 and the axes may not be perfectly perpendicular to each other. We assume that the magnitude of lens error is proportional to the distance from principal point instead of the camera coordinate origin. Therefore, theoretical camera coordinate [u-u_0,v_0,0] may be actually measured as:

$$\tilde{u} - u_0 = a(u-u_0) + b(v-v_0) - \delta u \tag{B4}$$

$$\tilde{v} - v_0 = c(v-v_0) - \delta v \tag{B5}$$

In the matrix form equation B4 and B5 can be combined as:

$$\begin{bmatrix} \tilde{u}-u_0 \\ \tilde{v}-v_0 \\ -f \end{bmatrix} = \begin{bmatrix} a & b & \delta u/f \\ 0 & c & \delta v/f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u-u_0 \\ v-v_0 \\ -f \end{bmatrix} \tag{B6}$$

Combining equation B3 and B6, we get

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ -f \end{bmatrix} = m \begin{bmatrix} a & b & \delta u/f \\ 0 & c & \delta v/f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \\ 0 \end{bmatrix} \tag{B7}$$

Compensations for images δu, and δu can be modeled as even power polynomials to secure rotational symmetry. It is common to use the polynomials of degree up to six.

$$\delta u = (\tilde{u} - u_0)(k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{B8}$$

$$\delta v = (\tilde{v} - v_0)(k_1 r^2 + k_2 r^4 + k_3 r^6) \tag{B9}$$

Where, $r^2 = (\tilde{u} - \tilde{u}_0)^2 + (\tilde{v} - \tilde{v}_0)^2$. Physical camera parameters are commonly divided into extrinsic and intrinsic parameters. In equation B7, $r_{11}.....r_{33}, x_0, y_0 \; and \; z_0$, total of nine parameters are called extrinsic camera parameters. Similarly, $a, b, c, u_0, v_0, k_1, k_2, k_3, f$ are called intrinsic camera parameters. Extrinsic parameters are needed to transform object coordinates to a camera centered coordinate frame. The purpose of the camera calibration is to determine the value of intrinsic and extrinsic camera parameters from many known scene points, *x, y* and *z*.

Epipolar constraint

Epipolar arises from the geometry of stereo vision. A point in the scene X is projected into point x1 on R1 plane, As we know that any point along the line x1c1 would have been projected into the same point x1. But, the points along the line x1C1 would be projected along the line E2x2 in the image plane R2. Therefore given a point x1, all the possible matches in plane R2 lies along the line ep2 called epipolar line of point x1. As all the rays forming image on plane R1 passes through lens center C1, all the epipolar lines passes from epipole E2 which is the projection point C1 onto image plane R2. Another point of epipole ep2 can be found by finding the infinite point along the line x1C1 and then projecting onto plane R2.

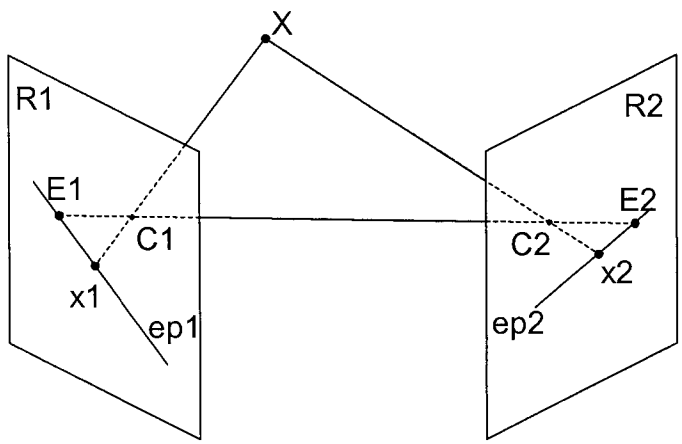


Figure B2. Depiction of epipolar constraint, ep1 and ep2 are two epipolar lines for left and right images

In this research the planes R1 and R2 are kept parallel. When R1 and R2 are parallel, Ea and E2 are at infinity and hence the epipolar lines are parallel. This makes it easier to search for a fracture object in a pair of the images.